


```

QY 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112
DB 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112

RESULT 2
US-09-062-465-1
Sequence 1, Application US/09062465
Patent No. 6465422
GENERAL INFORMATION:
APPLICANT: Schmidt, Ann Marie
APPLICANT: Stern, David
TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
FILE REFERENCE: 55424
CURRENT APPLICATION NUMBER: US/09/062465
CURRENT FILING DATE: 1998 04 17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 442
TYPE: PRT
ORGANISM: Human
US-09-062-465-1

Query Match
Best Local Similarity 100.0%; Score 595; DB 4; Length 442;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARIGEPVLKTCGAPKPPQPGKLNIGRTFAWKVLSPOGGGWDVSVARVLPG 60
DB 1 AONTARIGEPVLKTCGAPKPPQPGKLNIGRTFAWKVLSPOGGGWDVSVARVLPG 60

QY 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112
DB 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112

RESULT 3
US-08-643-148-2
Sequence 2, Application US/08643148
Patent No. 5864018
GENERAL INFORMATION:
APPLICANT: MORSE, MICHAEL J.
APPLICANT: NAGASHIMA, MARIKO
APPLICANT: HOLLANDER, DORIS A.
TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
FILE REFERENCE: 24
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: TOWNSEND & TOWNSEND & CREW LLP
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,148
FILING DATE: 16 APR 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY ESQ., MATTHEW B.
REGISTRATION NUMBER: 49,787
REFERENCE/DOCKET NUMBER: 014618-00560005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422

```

```

INFORMATION FOR SEQ ID NO: 2;
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-643-148-2

Query Match
Best Local Similarity 100.0%; Score 595; DB 2; Length 440;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARIGEPVLKTCGAPKPPQPGKLNIGRTFAWKVLSPOGGGWDVSVARVLPG 60
DB 1 AONTARIGEPVLKTCGAPKPPQPGKLNIGRTFAWKVLSPOGGGWDVSVARVLPG 60

QY 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112
DB 61 SLFLPVGIGDEGIFRCQAMNNGKTKSNYKRVVYQIPGKPEIVDSASELT 112

RESULT 4
US-08-442-016-5
Sequence 5, Application US/08442016
Patent No. 5968768
GENERAL INFORMATION:
APPLICANT: HAYNES, MARTIN F.
APPLICANT: ARUFFO, ALEJANDRO
APPLICANT: PATEL, DHAVALKUMAR
APPLICANT: BOWEN, MICHAEL A.
APPLICANT: MARQUARDT, HANS
TITLE OF INVENTION: CD6 LIGAND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,016
FILING DATE: 01 MAY 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08/443,450
FILING DATE: 02 NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08/143,904
FILING DATE: 02 NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 42,655
REFERENCE/DOCKET NUMBER: 1579 95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5;
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-442-016-5

Query Match
Score 558; DB 2; Length 278;

```

Best Local Similarity 99.98, Pred. No. 2,le 59,
Matches 104; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 67

Db 1 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 60

QY 68 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPKPEIVDSASELT 112

Db 61 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPKPEIVDSASELT 105

RESULT 5

US-08-684-594-5

; Sequence 5, Application US/08684594

; Patent No. 5998172

; GENERAL INFORMATION:

; APPLICANT: HAYNES, BARTON F.

; APPLICANT: ARUFFO, ALFANDRO

; APPLICANT: PATEL, DHAVALKUMAR

; APPLICANT: HOWEN, MICHAEL A.

; APPLICANT: MARQUARDT, HANS

; TITLE OF INVENTION: CD6 LIGAND

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHIVE P

; STREET: 1100 NORTH CLARE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/084,594

; FILING DATE: 18-JUL-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/432,016

; FILING DATE: 01-MAY-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/333,350

; FILING DATE: 02-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/143,903

; FILING DATE: 02-NOV-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: WILSON, MARY J

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 1579-112

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816 4100

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 278 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-684-594-5

Query Match

Best Local Similarity 99.98, Score 558, DB 2, Length 278,

Matches 104; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 67

Db 1 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 60

QY 68 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPKPEIVDSASELT 112

Db 61 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPKPEIVDSASELT 105

RESULT 6

US-09-062-365-5

; Sequence 5, Application US/09062365

; Patent No. 6465422

; GENERAL INFORMATION:

; APPLICANT: Schmidt, Ann Marie

; APPLICANT: Stern, David

; TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A

; TITLE OF INVENTION: SUBJECT

; FILE REFERENCE: 55424

; CURRENT APPLICATION NUMBER: US/09/062,365

; CURRENT FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 5

; LENGTH: 30

; TYPE: PRT

; ORGANISM: Human

US-09-062-365-5

Query Match

Best Local Similarity 100.0%; Pred. No. 6,3e-13; Length 30,

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQNITAIGEPVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 67

Db 1 AQNITAIGEPVLKCKGAPKKPQRLNIGRTIPANKVLSQGGGPDWSVARVLPNGSLFLPAV 60

RESULT 7

US-09-041-886-25

; Sequence 25, Application US/09041886

; Patent No. 6235872

; GENERAL INFORMATION:

; APPLICANT: Bredeson, Dale E.

; APPLICANT: Rabiadeh, Shantoz

; TITLE OF INVENTION: Peptide Mimetic Peptides, Dependence

; TITLE OF INVENTION: Polypeptides and Methods of Use

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell & Flores LLP

; STREET: 4370 La Jolla Village drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/041,886

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-1J 2626

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1447 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

us-09-872-185b-1.rai

Fri May 30 17:16:31 2003

```

US-09-041-886-25
Query Match
Best Local Similarity 41.4% Score 114; DB 4; Length 1447;
Matches 31; Conservative 20; Mismatches 38; Indels 10; Gaps 3;

QY 2 QNITAKGEPLVLCCKGAPKKPPQRIEHWKLNFGHTFAWKVLSPQGGPWVSARVLPNGS 61
DB 147 ESVTAFMGDTVLKCE-VIGEPMTTHWQNKQ-----QDLTPI---PCDSRVVVLPSGA 196

QY 62 LFLPVAVGIDQEGIFRCQAMNNGKETKSNRYRVVYVLPNG 100
DB 197 LQISRLQPGDGLGYRCSARNPASSRTGNEAEVRILSDPG 235

RESULT 8
PCT-US94-05277-2
Sequence 2, Application US/08752307H
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Discrete
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08752, 307H
FILING DATE: 19 NOV-1996
CLASSIFICATION: 445
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: McKie & Beckett, P.C., Anita L.
REGISTRATION NUMBER: 35, 283
REFERENCE/DOCKET NUMBER: 09404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-8906
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08 752-307H-12
Query Match 16.3% Score 97; DB 2; Length 607;
Best Local Similarity 40.5% Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICRPLVLCCKGAPKKPPQRIEHWKLNFGHTFAWKVLSPQGGPWVSARVLPNGSLEUP 65
DB 251 ALVGGQVTLCE-FAEGNEVPRIRKRVYDGS -LSPQ WTTA EPTLOIP 294

QY 66 AVGIQDEGIFRCQAMNNGKET 87
DB 295 SVSFEDEGTYRCEAFNSKGRH 416

RESULT 10
US-09-707-802-12
Sequence 12, Application US/09707802
Patent No. 6391586
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
ENCODING NOVEL SECRETED OR MEMBRANE ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston

```



```

? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? FILING DATE: 07-NOV-1996
? FILING DATE: 07-NOV-1996
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/752,307
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REFERENCE/DOCKET NUMBER: 35,283
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-542-5070
? TELEFAX: 617-542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 607 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-707-802-12

Query Match 16.3%, Score 97, DB 4; Length 607;
Best Local Similarity 30.5%, Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICEP-VIKCKGA-KKPKGLEWKNIGKTEAKKVLSPGSGPWSVAKVLPNSGLFLP 65
DB 251 ALVQQVTLSC-PFQNPVPRIKRWKVDGS-----LSQ-----WTTA-----EPTLOIP 294

QY 66 AVGIQDGGIFPCQAMNNGKET 87
DB 295 SVSFEDEGTVECAENSKGDEI 316

RESULT 11
US-09-991-326-12
? Sequence 12, Application US/09091326
? Patent No. 6395872
? GENERAL INFORMATION:
? APPLICANT: McCarthy, Sean A.
? CO-INVENTOR: David P.
? CO-INVENTOR: Douglas A.
? TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
? ENCODING NUCLEOTIDE SEQUENCE OF MEMBRANE-ASSOCIATED PROTEIN
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson, P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? FILING DATE: 07-NOV-1996
? FILING DATE: 07-NOV-1996
? APPLICATION NUMBER: US/09/991,326
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/752,307

```

```

? FILING DATE: 19-NOV-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REGISTRATION NUMBER: 35,283
? REFERENCE/DOCKET NUMBER: 09404/020002
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-542-5070
? TELEFAX: 617-542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 607 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-991-326-12

Query Match 16.3%, Score 97, DB 4; Length 607;
Best Local Similarity 30.5%, Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICEP-VIKCKGA-KKPKGLEWKNIGKTEAKKVLSPGSGPWSVAKVLPNSGLFLP 65
DB 251 ALVQQVTLSC-PFQNPVPRIKRWKVDGS-----LSQ-----WTTA-----EPTLOIP 294

QY 66 AVGIQDGGIFPCQAMNNGKET 87
DB 295 SVSFEDEGTVECAENSKGDEI 316

RESULT 12
US-08-506-296B-14
? Sequence 14, Application US/08506296B
? Patent No. 6313265
? GENERAL INFORMATION:
? APPLICANT: Wallis, Greg
? APPLICANT: Cunningham, Bruce A.
? APPLICANT: Crossin, Kathryn L.
? TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
? TITLE OF INVENTION: CONTAINING FIBROBLAST TYPE III REPEATS AND METHODS OF USE
? NUMBER OF SEQUENCES: 77
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: The Scripps Research Institute
? STREET: 10550 No. 631265th Torrey Pines Road, TPC-8
? CITY: La Jolla
? STATE: California
? COUNTRY: U.S.
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/506,296B
? FILING DATE: 24-JUL-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Filling, Thomas
? REGISTRATION NUMBER: 34,153
? REFERENCE/DOCKET NUMBER: TSRI 488.0
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 554-2937
? TELEFAX: (619) 554-6312
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1253 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-506-296B-14

```

```

Query Match 15.2% Score 96.5% EB 47 Length 1253
Best Local Similarity 32.4% pred. No. 0.0085
Matches 35; Conservative 127 Mismatches 40; Indels 21; Gaps 4

QY 1 AONTIATGCEPLVKN KAPAKPQKLEKWLNTGCTPAKVLSTQGGCPWLSVARVLP 58
Db 249 SSNVLALQGLVLECTAGEP-----TPTIKWLKPSGMPALRVTYQNH 293
QY 59 NSLFLPVGIGDHGTGTCQAMNNGKRETKSNYKRVYQIP --GKPE 103
Db 244 NKTLQKVGEDHDEYHCTAENSIG-SARHAYVYVTEAAPYLHKPQ 340

RESULT 13
US 08 441-844B 18
Sequence 18, Application US/08441844B
Patent No. 5872225
GENERAL INFORMATION:
APPLICANT: Lemmon, Vance
TITLE OF INVENTION: A Method for Characterizing the
TITLE OF INVENTION: Nucleotide Sequence of L10AM and
Patent No. 5872225
TITLE OF INVENTION: the Nucleotide Sequence
TITLE OF INVENTION: Characterized Thereby
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSER: Fay, Sharpe, Beall, Fagan,
ADDRESSER: Minnich & McKee
STREET: 1100 Superior Avenue
STREET: Suite 700
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2518
COMPILEX READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
MEDIUM TYPE: storable
COMPUTER: Compaq Prolinea 5100e
OPERATING SYSTEM: DOS 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,843B
FILING DATE: No. 5872225, earlier 18, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 29, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,174
REFERENCE/DOCKET NUMBER: CWR 2 149 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861 5582
TELEFAX: (216) 241 1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 96
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: irrelevant
ANIT SENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
INDIVIDUAL ISOLATE: 17 18 week fetus
IMMEDIATE SOURCE:
LIBRARY: Stratagene cDNA Library 936206
CLONE: synthesis of 4 clones
PUBLICATION INFORMATION:
AUTHORS: Blavie, Mary Louise
AUTHORS: Lemmon, Vance

```

```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: 17-18 week fetus
; IMMEDIATE SOURCE:
; LIBRARY: Stratagene cDNA Library 936206
; CLONE: synthesis of 4 clones
; PUBLICATION INFORMATION:
; AUTHORS: Ilavina, Mary Louise
; AUTHORS: Lemmon, Vance
; TITLE: Molecular structure and functional
; TITLE: testing of human IL6M: an interspecies
; TITLE: comparison.
; JOURNAL: GENOMICS
; VOLUME: 11
; ISSUE:
; PAGES: 416-423
; DATE: 1991
; US-08-427-497E-23

```

```

Query Match: 16.1%; Score 95.5; DB 2; Length 96;
Best Local Similarity 42.7%; Pred. No. 0.0029;
Matches 33; Conservative 11; Mismatches 39; Indels 18; Gaps 3;

QY 1 AONITARIHGEPLVLC--KGAPKPPQRLKWLNTGRTKAWKVLSPGGGWDVSVARVLP 58
Db 12 SSHLVALQGPVLECIAGFP .....TPTIKWLPSGPMFADRVTYQNH 56
QY 59 NGSILFPAVGIDEGIFRCQAMNRNGKETKSNRYVQIP 99
Db 57 NKTILQLLKVGEDDDGEYRCIAENSLG-SARHAYVTVLEAAP 96

```

```

Search completed: May 30, 2003, 15.59.50
Job time: 10.0811 secs

```

```

; LENGTH: 611 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-752-307B-10

```

```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: 17-18 week fetus
; IMMEDIATE SOURCE:
; LIBRARY: Stratagene cDNA Library 936206
; CLONE: synthesis of 4 clones
; PUBLICATION INFORMATION:
; AUTHORS: Ilavina, Mary Louise
; AUTHORS: Lemmon, Vance
; TITLE: Molecular structure and functional
; TITLE: testing of human IL6M: an interspecies
; TITLE: comparison.
; JOURNAL: GENOMICS
; VOLUME: 11
; ISSUE:
; PAGES: 416-423
; DATE: 1991
; US-08-427-497E-23

```

```

Query Match: 16.1%; Score 95; DB 2; Length 96;
Best Local Similarity 42.7%; Pred. No. 0.0029;
Matches 33; Conservative 11; Mismatches 39; Indels 18; Gaps 3;

QY 1 AONITARIHGEPLVLC--KGAPKPPQRLKWLNTGRTKAWKVLSPGGGWDVSVARVLP 58
Db 12 SSHLVALQGPVLECIAGFP .....TPTIKWLPSGPMFADRVTYQNH 56
QY 59 NGSILFPAVGIDEGIFRCQAMNRNGKETKSNRYVQIP 99
Db 57 NKTILQLLKVGEDDDGEYRCIAENSLG-SARHAYVTVLEAAP 96

```

```

RESULT 15
US-08-752-307B-10
Sequence 10, Application US/08752307B
Patent No. 5952171
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
NUMBER OF INVENTIONS: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.307B
FILING DATE: 19-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejoh, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

```



GenCore version 5.1.6
Copyright (c) 1993-2003 Compugen Ltd.

OM protein - protein search using SW model

Run on: May 30, 2003, 14:53:06; Search time 35.4775 seconds
(without alignments)
585.775 Million cell updates/sec

Title: US-09-872-185b-1
Perfect score: 595
Sequence: 1 AGCTTATCTGCTVTKKCAP VPVVGPERETVISASFLI 112

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Search: 908170 seqs, 13275620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 98
Maximum Match 1008
Listing first 45 summaries

Database: A_Genescat_10102			
	1	2	3
1:	/SID22/gqdata/geneseq/geneseq-emb1/AA1980.DAT		
2:	/SID22/gqdata/geneseq/geneseq-emb1/AA1981.DAT		
3:	/SID22/gqdata/geneseq/geneseq-emb1/AA1982.DAT		
4:	/SID22/gqdata/geneseq/geneseq-emb1/AA1983.DAT		
5:	/SID22/gqdata/geneseq/geneseq-emb1/AA1984.DAT		
6:	/SID22/gqdata/geneseq/geneseq-emb1/AA1985.DAT		
7:	/SID22/gqdata/geneseq/geneseq-emb1/AA1986.DAT		
8:	/SID22/gqdata/geneseq/geneseq-emb1/AA1987.DAT		
9:	/SID22/gqdata/geneseq/geneseq-emb1/AA1988.DAT		
10:	/SID22/gqdata/geneseq/geneseq-emb1/AA1989.DAT		
11:	/SID22/gqdata/geneseq/geneseq-emb1/AA1990.DAT		
12:	/SID22/gqdata/geneseq/geneseq-emb1/AA1991.DAT		
13:	/SID22/gqdata/geneseq/geneseq-emb1/AA1992.DAT		
14:	/SID22/gqdata/geneseq/geneseq-emb1/AA1993.DAT		
15:	/SID22/gqdata/geneseq/geneseq-emb1/AA1994.DAT		
16:	/SID22/gqdata/geneseq/geneseq-emb1/AA1995.DAT		
17:	/SID22/gqdata/geneseq/geneseq-emb1/AA1996.DAT		
18:	/SID22/gqdata/geneseq/geneseq-emb1/AA1997.DAT		
19:	/SID22/gqdata/geneseq/geneseq-emb1/AA1998.DAT		
20:	/SID22/gqdata/geneseq/geneseq-emb1/AA1999.DAT		
21:	/SID22/gqdata/geneseq/geneseq-emb1/AA2000.DAT		
22:	/SID22/gqdata/geneseq/geneseq-emb1/AA2001.DAT		
23:	/SID22/gqdata/geneseq/geneseq-emb1/AA2002.DAT		

pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	595	100.0	112	23	AAU77543
2	595	100.0	416	23	AAE23218
3	595	100.0	318	18	AAW33754
4	595	100.0	332	21	AAV52130
5	595	100.0	339	23	AAW48746
6	595	100.0	340	18	AAW44199
7	595	100.0	340	18	AAW33753
8	595	100.0	404	22	AAW33753
9	595	100.0	404	23	AAW48746
10	591	99.3	404	23	AAU77543

11	591	99.3	404	23	AAU77543	Human receptor for
12	513.5	86.3	416	23	AAE23218	Cow receptor for a
13	513.5	86.3	416	23	AAU77543	Bovine receptor for
14	512.5	86.1	403	23	AAE23200	Murine receptor for
15	512.5	86.1	403	23	AAU77544	Murine receptor for
16	162	27.2	30	20	AAU09449	Human RAGE V-domain
17	162	27.2	30	21	AAU52134	Human receptor to
18	162	27.2	30	23	AAW48746	Human RAGE V-domain
19	149	25.0	30	20	AAU09350	Mouse RAGE V-domain
20	145	24.4	30	20	AAU09352	Bovine RAGE V-domain
21	133	22.4	30	20	AAU09351	Pat RAGE V domain
22	127	21.3	30	18	AAW23337	N-terminal sequence
23	114	19.2	1417	16	AAW48753	Deleted in colorec
24	114	19.2	1417	20	AAU33498	Human DCC protein
25	114	19.2	1447	22	AAW50693	Human UNC-40 prote
26	114	19.2	1728	12	AAU11144	Deleted in colorec
27	104	17.5	2587	23	AAU47933	Human OCP. Homo s
28	101	17.0	2016	22	AAH66424	Deleted in colorec
29	99	16.6	479	22	AAW48752	Protophila melanog
30	99	16.6	627	23	AAH34079	Amino acid sequenc
31	99	16.6	628	22	AAW5805	Human pro peptide
32	99	16.6	628	22	AAW44469	Human leucine-rich
33	99	16.6	628	22	AAW47523	Amino acid sequenc
34	99	16.6	636	23	AAU32870	Amino acid sequenc
35	96.5	16.2	1253	23	AAU10648	Novel human secret
36	95.6	15.2	1257	20	AAW41152	Human L1 cell adhe
37	95	16.0	782	23	AAW53278	Human polypeptide
38	95	16.0	844	23	AAW53277	Novel human diago
39	94.5	15.9	898	22	AAU12152	Novel human diago
40	94	15.8	1247	22	AAW56337	Novel protein kin
41	94	15.8	1289	20	AAU27163	Peptide Seq ID No.
42	94	15.8	1289	21	AAW56781	Human trad protein
43	93.5	15.7	2993	22	AAW23884	Novel human diago
44	93.5	15.7	5266	22	AAW08561	Novel human diago
45	93	15.6	222	23	AAU83128	Novel secreted pro

ALIGNMENTS

RESULT 1	AAU48747	AAU48747 standard, proteins, 112 AA.
XX	AAU48747;	
XX	AAU48747;	
XX	02 APP 2002 (first entry)	
XX	Human RAGE protein V-domain SEQ ID NO 3.	
XX	Human RAGE receptor for advanced glycosylated endproduct; receptor;	
XX	anti-diabetic; neuroprotective; cytoprotective; anti-inflammatory; vasculoprotective;	
XX	keratinocyte; dermatological; anti-atherosclerotic; osteotropic; diabetes;	
XX	Alzheimer's disease; cancer; inflammation; kidney failure;	
XX	systemic lupus; nephritis; erectile dysfunction; atherosclerosis.	
OS	Homo sapiens.	
XX	W0200192892 A2.	
XX	06-DEC-2001.	
XX	30-MAY-2001; 2001WO-0517447	
XX	26-MAY-2000; 2000US-257312P.	
XX	05-MAR-2001; 2001US-0799152.	
XX	(TRAN-) TRANS TECH PHARMA.	
XX	Shatbar M;	
XX	WPI: 2002 114372/15.	

PT Detecting a receptor for advanced glycosylated end-products (AGE)
 PT modulators, for treating e.g., cancer, diabetes or inflammation,
 PT comprises measuring the amount of bound anti-AGE antibody
 XX
 PS Claim 4; Fig 2; 49pp; English.
 XX
 CC The invention relates to detection receptor for advanced glycosylated
 CC end-products (PAGE) modulators comprises determining the amount of PAGE
 CC protein or its fragment bound to the pre-adsorbed liquid by measuring the
 CC amount of anti-AGE antibody bound to the solid surface, the method is
 CC useful for rapid, high-throughput identification of compounds that
 CC modulate AGE. The compounds are useful for treating symptoms of diabetes
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis
 CC or inflammatory lupus nephritis, erectile dysfunction and
 CC atherosclerosis.
 XX
 SS Sequence 112 AA:
 Query Match 100.0%; Score 595; DB 24; Length 112;
 Best Local Similarity 100.0%; Pred. No. 5; 9c 5%;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AUNTARIGEPVLKCKAKPKPPQRLWKLNTGTEAWKVLSPQGGPWDSVARVLPNG 60
 ID 1 AUNTARIGEPVLKCKAKPKPPQRLWKLNTGTEAWKVLSPQGGPWDSVARVLPNG 60
 QY 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112
 ID 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112
 RESULT 2
 AAW44200
 ID AAW44200 standard; Protein: 418 AA.
 AC AAW44200;
 XX
 XX 14 MAY 1998 (first entry)
 DE Human mature receptor to an advanced glycosylation end product.
 KW Human; soluble receptor; advanced glycosylation end product; AGE;
 KW AGE; antibody; vascular permeability; diabetes mellitus.
 XX Homo sapiens.
 ES
 FH Key Location/Qualifiers
 FT Misc difference 66
 FT /note= "encoded by cct"
 XX
 UN W09799125-A1.
 XX
 PD 23 OCT 1997.
 XX
 PF 11 APR 1997; 97W0 EP01844.
 XX
 PR 16 APR 1996; 96US 0633148.
 XX
 FA (SCH) SCHERING PATENT AG;
 XX Hollander DA, Morser MJ, Nagashima M;
 DE WPI: 1997 568580/51.
 DR N PSDB: AAV12495.
 XX
 CC Anti advanced glycosylation end product polypeptide antibody
 PT prevents receptor binding and therefore reduces vascular
 PT permeability, useful to treat diabetes mellitus
 XX
 PS Claim 2; Page 42-44; 99pp; English.
 XX
 CC The present sequence represents a mature human receptor to an advanced

CC glycosylation end product (AGE) polypeptide. The present invention
 CC describes an isolated antibody (Ab), specifically immunoreactive with
 CC AGE. Advanced glycosylation end products (AGE) of proteins are
 CC non-enzymatically glycosylated proteins, which accumulate in vascular
 CC tissue in aging, and at an accelerated rate in individuals with
 CC diabetes. The Ab, which prevents the interaction between AGE and its
 CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
 CC treat diabetes mellitus symptoms, e.g., microvascular pathology, occlusive
 CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
 CC associated amyloidosis or atherosclerosis. The Ab can also be used for
 CC the isolation and purification of human RAGE polypeptide.
 XX
 SS Sequence 318 AA:
 Query Match 100.0%; Score 595; DB 18; Length 318;
 Best Local Similarity 100.0%; Pred. No. 26; 55;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AUNTARIGEPVLKCKAKPKPPQRLWKLNTGTEAWKVLSPQGGPWDSVARVLPNG 60
 ID 1 AUNTARIGEPVLKCKAKPKPPQRLWKLNTGTEAWKVLSPQGGPWDSVARVLPNG 60
 QY 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112
 ID 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112
 RESULT 4
 AAW44754
 ID AAW44754 standard; Protein: 418 AA.
 AC AAW44754;
 XX
 XX 08-MAY 1998 (first entry)
 XX
 DE Human RAGE polypeptide (418 amino acid residues).
 KW Advanced glycosylation end-product receptor; RAGE; screening; AGE;
 KW vascular permeability; diabetes mellitus; treatment; atherosclerosis;
 KW Alzheimer's disease.
 XX Homo sapiens.
 ES
 FH W09799121-A1.
 XX
 PD 23 OCT 1997.
 XX
 PF 11-APR 1997; 97W0 EP-01844.
 XX
 PR 16-APR 1996; 96US 0633147.
 XX (SCH) SCHERING AG.
 XX Morser MJ, Nagashima M;
 XX WPI: 1997 526458/48.
 DR N PSDB: AAV06518.
 XX
 CC New soluble advanced glycosylation end product receptor polypeptide
 PT used for reducing vascular permeability, complications of diabetes
 PT etc., also for purification and to screen for modulators
 XX
 PS Claim 4; Fig 1b; 91pp; English.
 XX
 CC This is a human advanced glycosylation end product receptor (PAGE)
 CC polypeptide (418 amino acid residues). The PAGE polypeptides and its
 CC active fragments or their mimetics, inhibit interaction between advanced
 CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They
 CC are used to treat diseases associated with AGE/RAGE interaction, such as
 CC increased vascular permeability, diabetes mellitus (particularly
 CC complications such as micro or macro vasculopathy or occlusive vascular
 CC disorders such as neuropathy, nephropathy, retinopathy or
 CC atherosclerosis) or haemodialysis associated amyloidosis, also activation

CC of microglial cells by beta amyloid peptides in Alzheimer's disease of
 CC age related disorders such as oxidative stress. These RAGE polypeptides
 CC are also used, when immobilised, to purify AGE from a protein mixture and
 CC to screen for compounds that are agonists and antagonists of AGE/RAGE
 CC interaction. They can also be used diagnostically to detect abnormal
 CC levels of AGE. Antibodies against RAGE polypeptides are useful as
 CC immunosay reagents for measurement of RAGE levels, and as inhibitors of
 CC interaction between AGE and RAGE or other receptors and for purification
 CC and quantification of AGE polypeptides. The following table lists the
 CC used to express recombinant RAGE and as probes for isolating related
 CC genes.

XX Sequence 318 AA;

Query Match 100.0%; Score 535; DB 18; Length 318;
 Best Local Similarity 100.0%; Pred. No 2e-55;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTATIGEPVLKCKGAPKPPORLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60
 DB 1 AONTATIGEPVLKCKGAPKPPORLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60
 QY 61 SLFLPVGIGDEGIFRQAMNNGKTKSNYPVPVQIPGKPEIVDSASELT 112
 DB 61 SLFLPVGIGDEGIFRQAMNNGKTKSNYPVPVQIPGKPEIVDSASELT 112

RESULT 4

AAV52130

ID AAV52130 standard; protein: 332 AA.

XX

AC AAV52130;

XX

DT 28-MAY-2003 (first entry)

XX

DE Human Receptor to AGE (RAGE) amino acid sequence.

XX

KW Soluble receptor for advanced glycation endproducts (RAGE) tumour
 invasion; metastasis; amphotericin; neuron; inhibit; therapy.

XX

OS Homo sapiens.

XX

PN W09554485-A1.

XX

PD 28-OCT-1999

XX

PF 16-APR-1999; qaw-n508427

XX

PR 17-APR-1998; 98US-0062365

XX

PA (UNCO) UNIV COLUMBIA NEW YORK.

XX

PI Schmidt AM, Stern D;

XX

DR WPI; 2000-013260/01.

XX

PT Inhibiting tumour invasion or spreading by administration of soluble

XX

PT receptor for advanced glycation endproducts -

XX

PS Disclosure: page 10-11, 88pp, English.

XX

CC This is the amino acid sequence of the human soluble receptor for
 CC Advanced Glycation Endproducts (RAGE). RAGE interacts with a range of
 CC physiologically and pathophysiologically relevant ligands when
 CC considering tumour invasion. In normal developing neurons RAGE
 CC colocalizes with amphoterin which is a matrix associated polypeptide.
 CC The expression of both RAGE and amphoterin decreases after birth, but
 CC both have increased expression in tumours. RAGE polypeptides
 CC AAV52132-Y52135 are used in the invention in a method for inhibiting
 CC tumour invasion and metastasis. The method involves inhibiting tumour
 CC invasion and metastasis via administration of a therapeutically effective
 CC amount of the pharmaceutical composition containing a RAGE polypeptide.
 CC The invention also relates to a method for evaluating the ability of an

CC agent to inhibit tumour invasion in a local cellular environment. RAGE
 CC can be administered to a patient in a pharmaceutical y acceptable
 CC carrier.

XX Sequence 332 AA;

Query Match 100.0%; Score 535; DB 21; Length 332;
 Best Local Similarity 100.0%; Pred. No 2.2e-55;
 Matches 112; Conservative 0; Mismatches 0; Indels 5; Gaps 0;

QY 1 AONTATIGEPVLKCKGAPKPPORLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60
 DB 1 AONTATIGEPVLKCKGAPKPPORLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60
 QY 61 SLFLPVGIGDEGIFRQAMNNGKTKSNYPVPVQIPGKPEIVDSASELT 112
 DB 61 SLFLPVGIGDEGIFRQAMNNGKTKSNYPVPVQIPGKPEIVDSASELT 112

RESULT 5

AAM48746

ID AAM48746 standard; protein: 339 AA.

XX

AC AAM48746;

XX

DT 02-APR-2002 (first entry)

XX

DE Human RAGE protein SEQ ID NO 2

XX

KW Human, RAGE, receptor for advanced glycation endproduct, receptor,
 antidiabetic, neuroprotective, cytostatic, antiinflammatory, vasotrophic,
 nephrotrophic, dermatologic, antiarteriosclerotic, neurotrophic, diabetes;
 Alzheimer's disease; cancer; inflammation; kidney failure;
 systemic lupus, nephritis, erectile dysfunction, atherosclerosis.

XX

OS Homo sapiens.

XX

PN W0200192892-A2.

XX

PD 06-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US17447.

XX

PR 30-MAY-2000; 2000US-207342P.

XX

PF 05-MAR-2001; 2001US-0799152

XX

PA (TEAN-) TRANS TECH PHARMA.

XX

PI Shahbar M;

XX

DR WPI; 2002-114372/15.

XX

PT Detecting a receptor for advanced glycation endproducts (RAGE)

XX

PT modulators, for treating e.g., cancer, diabetes or inflammation,

XX

PT comprises measuring the amount of bound anti-RAGE antibody -

XX

XX Claim 2; Fig 2; 49pp; English.

XX

CC The invention relates to detecting receptor for advanced glycation
 CC endproducts (RAGE) modulators comprising determining the amount of RAGE
 CC protein or its fragment bound to the pre-adsorbed ligand by measuring the
 CC amount of anti-RAGE antibody bound to the solid surface. The method is
 CC useful for rapid, high-throughput identification of compounds that
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis
 CC or inflammatory lupus nephritis, erectile dysfunction and
 CC atherosclerosis.

XX

SQ Sequence 339 AA;

XX

Query Match 100.0%;

XX

Best Local Similarity 100.0%; Score 535; DB 23; Length 339;

XX

Matches 112; Conservative 0; Mismatches 0; Indels 5; Gaps 0;

RESULT 8
AAH81925
ID AAH81925 standard, protein, 404 AA.
XX AC AAB81925;
XX DT 15-JUN-2001 (first entry)
XX DE Extracorporeal circulation material receptor protein.
XX KW Extracorporeal circulation; carbonyl stress product; receptor;
XX KW diabetes; vascular lesion; excretory dysfunction
XX OS Unidentified.
XX PN W0200118060-A1.
XX PD 15-MAR-2001.
XX PF 08-SEP-1999, 2000W03-010612.
XX PR 08-SEP-1999; 99JP-0254463.
XX PA (TORA) TOPAY IND INC.
XX PI Shimizu S, Kubota M, Akiyama H, Usui M;
XX WP 2001-2903:4/30.
XX Material for extracorporeal circulation, applicable in selective
PT elimination of diabetic complication factors such as carbonyl stress
PT products caused by abnormally promoted carbonyl stress from excretory
PT dysfunction in vascular lesions
XX PS Claim 1; Page 31-32; 36pp; Japanese.
XX CC The present invention describes a material for extracorporeal circulation
CC which is made from a water-insoluble carrier immobilized with a protein
CC having the sequence shown here. The materials of the invention, including
CC adsorbents, are for extracorporeal circulation, which are applicable in
CC the selective elimination of diabetic complication factors from a body
CC fluid, and are therefore useful in treating vascular lesions like
CC arteriosclerosis due to carbonyl stress products caused by abnormally
CC promoted carbonyl stress from excretory dysfunction.
XX SQ Sequence 404 AA;
Query Match 100.0%; Score 595; DB 22; Length 404;
Best Local Similarity 100.0%; Pred. No. 2.7e-55;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 60
Db 23 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 82
QY 61 SLFLPAVGIDEGIFRCQAMNRNGKETKSNRYRVYQIPCKPEIVDSASELT 112
Db 83 SLFLPAVGIDEGIFRCQAMNRNGKETKSNRYRVYQIPCKPEIVDSASELT 134
RESULT 9
AAM48745
ID AAM48745 standard; protein; 404 AA.
XX AC AAM48745;
XX DT 02-APR-2002 (first entry)
XX DE Human RAGE protein SEQ ID NO 1.
XX KW Human; RAGE; receptor for advanced glycosylated endproduct; receptor;

KW antidiabetic; neuroprotective; cytosstatic; antiinflammatory; vasotropic;
KW nephrotropic; dermatological; antiarteriosclerotic; nootropic; diabetes;
KW Alzheimer's disease; cancer; inflammation; kidney failure;
KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.
OS Homo sapiens.
XX PN W0200192892-A2.
XX PD 06-DEC-2001
XX PF 30-MAY-2001; 2001W0-US17447.
XX PP 30-MAY-2000; 2000US-307342P.
XX PP 05-MAR-2001; 2001US-0799152.
XX PA (TRAN-) TRANS TECH PHARMA.
XX PI Shahbaz M;
XX WP 2002-114372/15.
XX Detecting a receptor for advanced glycosylated endproducts (RAGE)
PT modulators, for treating e.g., cancer, diabetes or inflammation,
PT comprises measuring the amount of bound anti-PAGE antibody
XX Claim 1; Fig 2; 49pp; English.
XX CC The invention relates to detecting receptor for advanced glycosylated
CC endproducts (RAGE) modulators comprising determining the amount of RAGE
CC protein or its fragment bound to the pre-adsorbed ligand by measuring the
CC amount of anti-RAGE antibody bound to the solid surface. The method is
CC useful for rapid, high-throughput identification of compounds that
CC modulate RAGE. The compounds are useful for treating symptoms of diabetes
CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's
CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis
CC or inflammatory lupus nephritis, erectile dysfunction and
XX atherosclerosis.
XX SQ Sequence 404 AA;
Query Match 100.0%; Score 595; DB 23; Length 404;
Best Local Similarity 100.0%; Pred. No. 2.7e-55;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 60
Db 23 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 82
QY 61 SLFLPAVGIDEGIFRCQAMNRNGKETKSNRYRVYQIPCKPEIVDSASELT 112
Db 83 SLFLPAVGIDEGIFRCQAMNRNGKETKSNRYRVYQIPCKPEIVDSASELT 134
RESULT 10
AAE23219
ID AAE23219 standard; protein; 404 AA.
XX AC AAE23219;
XX DT 27-AUG-2002 (first entry)
XX DE Human receptor for advanced glycosylation end product (RAGE) protein.
XX KW Human; Receptor for advanced glycosylation end product; RAGE; cardiant;
XX KW tissue growth; neointimal formation; blood vessel; restenosis; diabetes;
XX KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;
XX KW transgenic animal; acute thrombotic stroke; venous thrombosis.
OS Homo sapiens.
XX PN W0200250889-A3.
XX


```

XX 13-OCT-2000; 2000US-0687528.
XX (UYCO ) UNIV COLUMBIA NEW YORK.
XX Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;
XX WPI: 2002-426260/45.
XX N-PSDB; AAD36951.
XX Inhibiting new tissue growth or neointimal formation in blood vessels
XX of subject suffering from diabetes, stroke and preventing restenosis,
XX comprises administering inhibitor of receptor for advanced glycation
XX product .
XX Disclosure: Page 14; 43pp; English.
XX The invention relates to a method for inhibiting new tissue growth or
XX neointimal formation in blood vessels in a subject that has experienced
XX blood vessel injury and preventing exaggerated restenosis in a diabetic
XX subject. The method comprises administering an inhibitor of receptor for
XX advanced glycation/glycosylation end product (RAGE), so as to inhibit new
XX tissue growth or neointimal formation in subject's blood vessels and
XX preventing restenosis in the subject. The method is useful for inhibiting
XX new tissue growth or neointimal formation in blood vessels in a subject
XX like non-human animal, a transgenic non-human animal or a human suffering
XX from diabetes, acute thrombotic stroke, venous thrombosis, unstable
XX angina, myocardial infarction, abrupt closure following angioplasty or
XX stent placement, or thrombosis as a result of peripheral vascular surgery
XX The method is also useful for preventing restenosis and for determining
XX whether a compound inhibits new tissue growth in a blood vessel in a
XX subject. The present sequence is cow receptor for advanced glycosylation
XX end product (RAGE) protein.
XX SQ Sequence 416 AA;
Query Match 86.3%; Score 513.5; DB 23; Length 416;
Best Local Similarity 89.1%; Pred. No. 1.5e-46;
Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
QY 2 QNITATIREPIVLKCKGAPKPPQPLEWKLNTRTEAWKVLSPGSGPWFVSAPVLPNGS 61
DB 24 QNITATIRCKPLVINCCKGAPKPPQPLEWKLNTRTEAWKVLSPQ-GDPWDSVARVLPNGS 82
QY 62 LFLPAVGIDEGIERCOAMNNGKTKSNRYRVVYQIPGKPEIVDSASEL 111
DB 83 LLLPAVGIDQETFPATSPSGKTKSNRYRVVYQIPGKPEIVDPASEL 132
RESULT 13
AAU77542
XX AAU77542 standard; Protein: 416 AA.
XX AAU77542;
XX 05-JUN-2002 (first entry)
XX Bovine receptor for advanced glycosylation end product (RAGE).
XX Receptor for advanced glycation end product, RAGE, receptor,
XX amyloid beta peptide, blood-brain barrier, neurovascular stress,
XX cerebral vasoconstriction, suppression, cerebral blood flow enhancer,
XX cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;
XX Alzheimer's disease; Down's syndrome, head trauma, stroke, bovine.
XX OS Bos taurus.
XX WO20021419-A1
XX 21 FEB 2002
XX 14 AUG 2003; 2001WO-US25416.

```

```

PR 14-AUG-2000; 2000US-0638648
XX (UYCO ) UNIV COLUMBIA NEW YORK.
XX Stern DM, Schmidt AM, Yan SD, Zlokovic B;
XX WPI: 2002-257610/30.
XX N-PSDB; ABK10855.
XX Ameliorating neurovascular stress and decreasing cerebral
XX vasoconstriction in subject suffering from chronic/acute cerebral
XX amyloid angiopathy, by administering inhibitor of receptor for advanced
XX glycation endproduct .
XX Disclosure: Page 13-14; 68pp; English.
XX The invention describes a method of ameliorating neurovascular stress,
XX and decreasing cerebral vasoconstriction in subject suffering from
XX chronic or acute cerebral amyloid angiopathy, comprising administering
XX an inhibitor (I) of receptor for advanced glycation end product (RAGE).
XX (I) inhibits transcytosis of amyloid beta peptides across blood-brain
XX barrier, thus decreasing cerebral vasoconstriction and increasing
XX cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
XX subject, decreasing cerebral vasoconstriction in a transgenic non-human
XX animal (preferably, transgenic mouse overexpressing mutant human amyloid
XX beta precursor protein) or a human, suffering from chronic or acute
XX cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
XX ameliorating neurovascular stress comprising cerebral amyloid angiopathy
XX in a subject, where the neurovascular stress is caused by Alzheimer's
XX disease, aging, Down's syndrome, head trauma or stroke. This is the
XX amino acid sequence of bovine receptor for advanced glycation end
XX product (RAGE) described in the invention.
XX SQ Sequence 416 AA;
Query Match 86.3%; Score 513.5; DB 23; Length 416;
Best Local Similarity 89.1%; Pred. No. 1.5e-46;
Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
QY 2 QNITATIREPIVLKCKGAPKPPQPLEWKLNTRTEAWKVLSPGSGPWFVSAPVLPNGS 61
DB 24 QNITATIRCKPLVINCCKGAPKPPQPLEWKLNTRTEAWKVLSPQ-GDPWDSVARVLPNGS 82
QY 62 LFLPAVGIDEGIERCOAMNNGKTKSNRYRVVYQIPGKPEIVDSASEL 111
DB 83 LLLPAVGIDQETFPATSPSGKTKSNRYRVVYQIPGKPEIVDPASEL 132
RESULT 14
AAE23220
XX AAE23220 standard; Protein: 403 AA.
XX AAE23220;
XX 27-AUG-2002 (first entry)
XX Mouse receptor for advanced glycosylation end product (RAGE) protein.
XX Mouse: Receptor for advanced glycosylation end product; RAGE; cardiac;
XX tissue growth, neointima; formation, blood vessel, restenosis; diabetes,
XX myocardial infarction; angioplasty; peripheral vascular surgery; angina;
XX transgenic animal, acute thrombotic stroke, venous thrombosis.
XX OS Mus musculus.
XX WO200230889 A2.
XX 18-APR-2002.
XX 12-OCT-2001, 2001WO-US32036.
XX 13 OCT-2000; 2000US-0687528.
XX

```

PA (UYCO) UNIV COLUMBIA NEW YORK.

XX Stern DM, Schmidt AM, Marso S, Topol EJ, Lincott AM;

XX WPT: 2002 426260/45;

DB N PSDB; AAD46954;

XX Inhibiting new tissue growth or neointimal formation in blood vessels

PT of subject suffering from diabetes, stroke and preventing restenosis,

PI comprises administering inhibitor of receptor for advanced glycation end

PI product

XX Disclosure: Page 18; 4 pp; English.

XX The invention relates to a method for inhibiting new tissue growth or

CC neointimal formation in blood vessels in a subject that has experienced

CC blood vessel injury and preventing exaggerated restenosis in a diabetic

CC subject. The method comprises administering an inhibitor of receptor for

CC advanced glycation/lysosylation end product (RAGE), so as to inhibit new

CC tissue growth or neointimal formation in subject's blood vessels and

CC preventing restenosis in the subject. The method is useful for inhibiting

CC new tissue growth or neointimal formation in blood vessels in a subject

CC like non-human animal, a transgenic non-human animal or a human suffering

CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable

CC angina, myocardial infarction, abrupt closure following angioplasty or

CC stent placement, or thrombosis as a result of peripheral vascular surgery

CC The method is also useful for preventing restenosis and for determining

CC whether a compound inhibits new tissue growth in a blood vessel in a

CC subject. The present sequence is mouse receptor for advanced

CC glycosylation end product (RAGE) protein.

XX Sequence 403 AA;

SO Query Match 86.1%; Score 512.5; DB 24; Length 403;

Best Local Similarity 87.4%; Prod. No. 146 46;

Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 UNITAPGEPLVLEKCKGAKKKPKQRLKWLNTGRTFAWKVLSFGGGGWSVARVLPNGS 61

DB 24 UNITAPGEPLVLSCKGAKKKPKQRLKWLNTGRTFAWKVLSFGGGGWSVARVLPNGS 82

QY 62 LFLPAVGIDGEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 112

DB 83 LLLPATGIVDEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 134

RESULT 15

AAU77544

XX AAU77544 standard; Protein: 403 AA.

XX AAU77544;

XX 05-JUN 2002 (first entry)

XX Marine receptor for advanced glycosylation end product (RAGE).

XX Receptor for advanced glycation end product; RAGE; receptor;

XX amyloid beta peptide; blood brain barrier; neurovascular stress;

XX cerebral vasoconstriction suppressor; cerebral blood flow enhancer;

XX cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;

XX Alzheimer's disease; Down's syndrome; head trauma; stroke; mouse.

XX Mus musculus.

XX W0200214519-A1.

XX 21 FEB 2002.

XX 14 AUG 2001; 2001W00825416.

XX 14-AUG 2000; 2000US 06-60448.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Stern DM, Schmidt AM, Yan SB, Zlokovic B;

XX WPT: 2002 252616/40.

DB N PSDB; AAK10857; AAK10858.

XX Ameliorating neurovascular stress and decreasing cerebral

PT vasoconstriction in subject suffering from chronic/cerebral

PI amyloid angiopathy, by administering inhibitor of receptor for advanced

PI glycation endproduct

XX Disclosure: Page 17 18; 68pp; English.

XX The invention describes a method of ameliorating neurovascular stress,

CC and decreasing cerebral vasoconstriction in subject suffering from

CC chronic or acute cerebral amyloid angiopathy, comprising administering

CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).

CC (I) inhibits transcytosis of amyloid beta peptides across blood brain

CC barrier, thus decreasing cerebral vasoconstriction and increasing

CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a

CC subject, decreasing cerebral vasoconstriction in a transgenic non human

CC animal (preferably, transgenic mouse overexpressing mutant human amyloid

CC beta precursor protein) or a human, suffering from chronic or acute

CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for

CC ameliorating neurovascular stress in subject suffering from amyloid angiopathy

CC in a subject, where the neurovascular stress is caused by Alzheimer's

CC disease, aging, Down's syndrome, head trauma or stroke. Thus is the

CC amino acid sequence of murine receptor for advanced glycation end

CC product (RAGE) described in the invention.

XX Sequence 403 AA;

SO Query Match 86.1%; Score 512.5; DB 24; Length 403;

Best Local Similarity 87.4%; Prod. No. 146 46;

Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 UNITAPGEPLVLEKCKGAKKKPKQRLKWLNTGRTFAWKVLSFGGGGWSVARVLPNGS 61

DB 24 UNITAPGEPLVLSCKGAKKKPKQRLKWLNTGRTFAWKVLSFGGGGWSVARVLPNGS 82

QY 62 LFLPAVGIDGEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 112

DB 83 LLLPATGIVDEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 134

Search Completed: May 30, 2003, 15:55:04

Job Time : 27.4775 secs

? TITLE OF INVENTION: A METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A SUBJECT
? FILE NUMBER: 0575/50159
? CURRENT APPLICATION NUMBER: US/09/872,185B
? CURRENT FILING DATE: 2001 05 06
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 1
? LENGTH: 432
? TYPE: PRT
? ORGANISM: Human
US-09-872-185b-1

Query Match
Best Local Similarity 100.0% Score 595; DB 9; Length 432;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 60
DB 1 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 60
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 112
DB 61 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 112

RESULT 4
US-09-872-185b-2
? Sequence 2, Application US/09/2185B
? Patent No. US20020132799A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Herold, Kevin
? APPLICANT: Yan, Shi Du
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Lamster, Ira
? TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/09/872,185B
? CURRENT FILING DATE: 2001 06 01
? NUMBER OF SEQ ID NOS: 16
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 2
? LENGTH: 432
? TYPE: PRT
? ORGANISM: Human
US-09-872-185b-2

Query Match
Best Local Similarity 100.0% Score 595; DB 10; Length 432;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 60
DB 1 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 60
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 112
DB 61 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 112

RESULT 4
US-09-872-185b-4
? Sequence 4, Application US/09/872,185
? Patent No. US20040059424A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Wu, Jun
? TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/09/872,185
? CURRENT FILING DATE: 1996 11 22
? NUMBER OF SEQ ID NOS: 4

? SOFTWARE: Patent in version 4.1
? SEQ ID NO 4
? LENGTH: 405
? TYPE: PRT
? ORGANISM: Human
US-08-755-235-4

Query Match
Best Local Similarity 99.1% Score 591; DB 1; Length 405;
Matches 111; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 60
DB 24 AONTATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 82
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 112
DB 83 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 144

RESULT 5
US-08-755-235-2
? Sequence 2, Application US/08/755,235
? Patent No. US20040059424A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Wu, Jun
? TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/08/755,235
? CURRENT FILING DATE: 1996-11-22
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 2
? LENGTH: 416
? TYPE: PRT
? ORGANISM: Bovine
US-08-755-235-2

Query Match
Best Local Similarity 86.4% Score 513.5; DB 1; Length 416;
Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
QY 2 ONITATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 61
DB 24 ONITATIGEPVLVK*KGAKPKPPQRIEKLNTGRTAEAKVLSPPGGGPPWDSVARVLPNG 82
QY 62 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 111
DB 83 SLFLPAVGIGDEGIFRCQAMNNGKTKSNYRVVYQIPKPEIVDSASELT 142

RESULT 5
US-08-948-131-1
? Sequence 1, Application US/08/948,131
? Patent No. US20010054357A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David
? APPLICANT: Yan, Shi Du
? APPLICANT: Schmidt, Ann Marie
? TITLE OF INVENTION: Human Hinding Site of Kato and Uses
? TITLE OF INVENTION: Thero
? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESS: Cooper & Dunham
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10046
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,131
FILING DATE: 09-OCT-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 53447
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212 391 0526
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-948-131-1

Query Match 27.2% Score 162; DB 8; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30
|||||
Db 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30

RESULT 7

US-09-851-071-5
Sequence 5, Application US/09851071
Patent No. US2002017750A1
GENERAL INFORMATION:
APPLICANT: Schmidt, Anne Marie
TITLE OF INVENTION: A METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A SUBJECT
FILE REFERENCE: 0575/55424 Z/TPW/SHS/MCM
CURRENT APPLICATION NUMBER: US/09/851,071
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 30
TYPE: PRT
ORGANISM: Human
US-09-851-071-5

Query Match 27.2% Score 162; DB 9; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30
|||||
Db 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30

RESULT 8

US-09-872-185B-3
Sequence 3, Application US/09872185B
Patent No. US2002012279A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Herold, Kevin
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
FILE REFERENCE: 0575/64080
CURRENT APPLICATION NUMBER: US/09/872,185B

CURRENT FILING DATE: 2001-06-01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 30
TYPE: PRT
ORGANISM: Human
US-09-872-185B-3

Query Match 27.2% Score 162; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30
|||||
Db 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30

RESULT 9

US-09-872-185B-7
Sequence 7, Application US/09872185B
Patent No. US2002012279A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Herold, Kevin
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
FILE REFERENCE: 0575/64080
CURRENT APPLICATION NUMBER: US/09/872,185B
CURRENT FILING DATE: 2001-06-01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 30
TYPE: PRT
ORGANISM: Human
US-09-872-185B-7

Query Match 27.2% Score 162; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30
|||||
Db 1 AQNITARIGPVLVKCKGAPKKPPQPLEWK 30

RESULT 10

US-08-948-131-2
Sequence 2, Application US/08948131
Patent No. US20010053357A1
GENERAL INFORMATION:
APPLICANT: Stern, David
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: Ligand Binding Site of Rage and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

us-09-872-185b-1.rapb

Fri May 30 17:16:31 2003

```

1  APPLICATION NUMBER:  US/08/948,131
2  FILING DATE: 09-OCT 1997
3  CLASSIFICATION: 530
4  ATTORNEY/AGENT INFORMATION:
5  NAME: White, John P
6  REGISTRATION NUMBER: 28,678
7  REFERENCE/DOCKET NUMBER: 5447
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE: 212-278-0400
10 TELEFAX: 212-391-0526
11 INFORMATION FOR SEQ ID NO: 2:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 30 amino acids
14 TYPE: amino acid
15 STRANDEDNESS: single
16 TOPOLOGY: linear
17 MOLECULE TYPE: peptide
18 US-08-948-131-2

Query Match      25.0%; Score 149; DB 8; Length 30;
Best Local Similarity 94.1%; Pred. No. 8 to 08;
Matches 27; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 QNITARIGEPVLVKGAKKPPQQLWK 30
    ||||| ||||| ||||| ||||| |||||
DB 2 QNITARIGEPVLVSKGAKKPPQQLWK 30

RESULT 11
US-09-872-185B-4
Sequence 4, Application US/09872185B
Patent No. US20020122799A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Herold, Kevan
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
APPLICANT: Lamster, Ira
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
FILE REFERENCE: 0575/64080
CURRENT APPLICATION NUMBER: US/09/872,185B
CURRENT FILING DATE: 2001-06-01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 30
TYPE: PRT
ORGANISM: Murine
US-09-872-185B-4

Query Match      25.0%; Score 149; DB 10; Length 30;
Best Local Similarity 94.1%; Pred. No. 8, 30-08;
Matches 27; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 QNITARIGEPVLVKGAKKPPQQLWK 30
    ||||| ||||| ||||| ||||| |||||
DB 2 QNITARIGEPVLVSKGAKKPPQQLWK 30

```

```

Query Match      24.4%; Score 145; DB 8; Length 30;
Best Local Similarity 89.7%; Pred. No. 20-07;
Matches 26; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 QNITARIGEPVLVKGAKKPPQQLWK 30
    ||||| ||||| ||||| ||||| |||||
DB 2 QNITARIGEPVLVNCKGAKKPPQQLWK 30

```

```

RESULT 13
US-09-872-185B-6
Sequence 6, Application US/09872185B
Patent No. US20020122799A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Herold, Kevan
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
APPLICANT: Lamster, Ira
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
FILE REFERENCE: 0575/64080
CURRENT APPLICATION NUMBER: US/09/872,185B
CURRENT FILING DATE: 2001-06-01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 30
TYPE: PRT
ORGANISM: bovine
US-09-872-185B-6

Query Match      24.4%; Score 145; DB 10; Length 30;
Best Local Similarity 89.7%; Pred. No. 20-07;
Matches 26; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 QNITARIGEPVLVKGAKKPPQQLWK 30
    ||||| ||||| ||||| ||||| |||||
DB 2 QNITARIGEPVLVNCKGAKKPPQQLWK 30

```

```

RESULT 14
US-08-948-131-3
Sequence 3, Application US/08948131
Patent No. US20010053357A1

```

```

RESULT 12
US-08-948-131-4
Sequence 4, Application US/08948131
Patent No. US20010053357A1
GENERAL INFORMATION:
APPLICANT: Stern, David
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
APPLICANT: Lamster, Ira
TITLE OF INVENTION: Ligand Binding Site of RAGE and Uses
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas

```


GENERAL INFORMATION:
 APPLICANT: Stern, David
 APPLICANT: Yan, Shi Du
 APPLICANT: Schmidt, Ann Marie
 TITLE OF INVENTION: Ligand Binding Site of Rage and Uses
 TITLE OF INVENTION: Thetrol
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: Cooper & Dunham
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/948,131
 FILING DATE: 09-OCT-1997
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 53447
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-278-0400
 TELEFAX: 212-391-0536
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 30 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-948-131-3

Query Match: 22.4%, Score 133, Ins 8, Length 30;
 Best Local Similarity: 82.8%, Pred. No. 2.8e-06;
 Matches: 24; Conservative: 2; Mismatches: 3; Indels: 0; Gaps: 0;

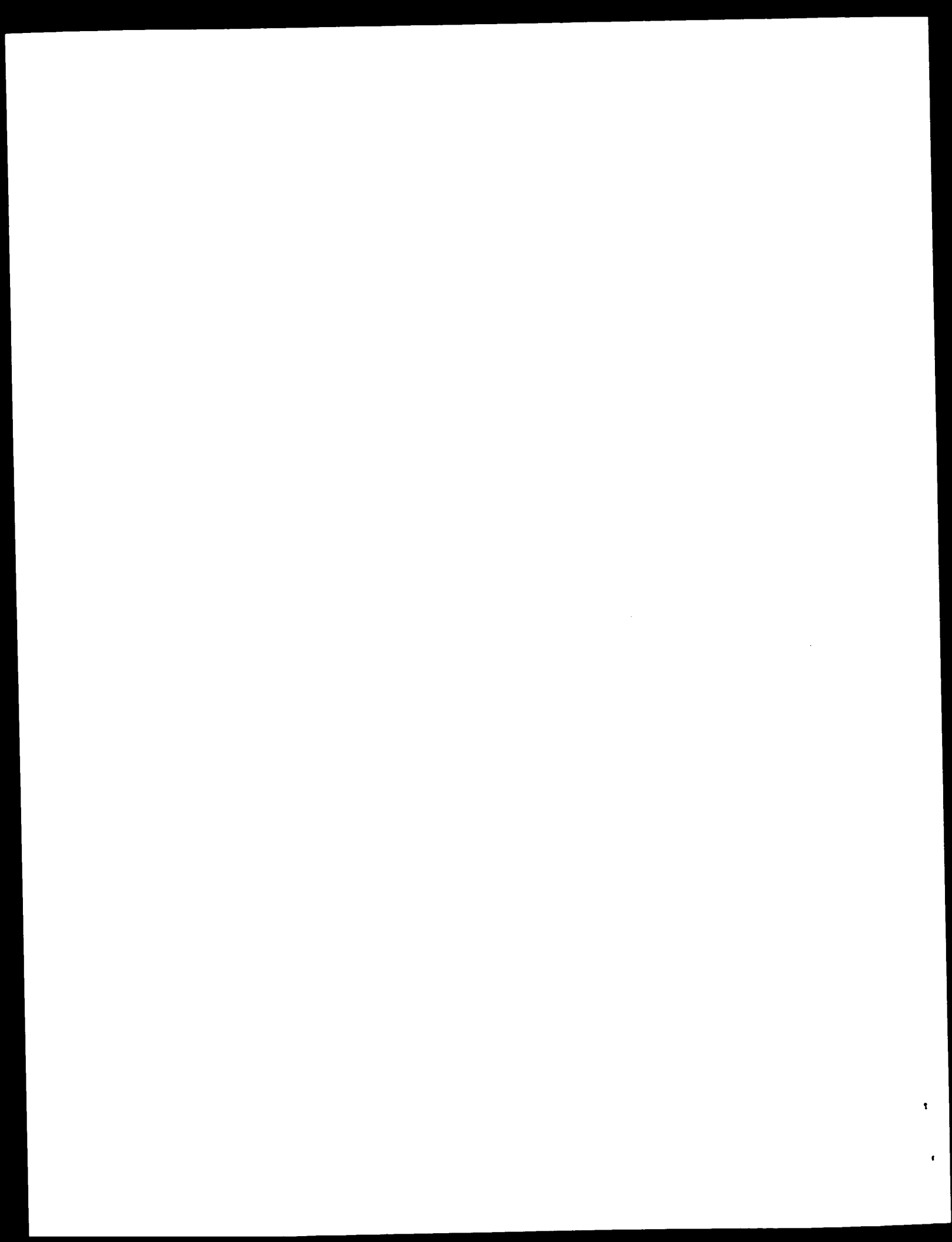
QY 2 QNITARIGEPVLKCKGAPKPKPPQRLWK 30
 DB 2 QNITARIGEPVLKCKGAPKPKPPQRLWK 30

RESULT 15
 US-09-872-185B-5
 Sequence 5, Application US/09872185B
 Patent No. US20020122799A1
 GENERAL INFORMATION:
 APPLICANT: Stern, David M.
 APPLICANT: Herold, Kevin
 APPLICANT: Yan, Shi Du
 APPLICANT: Schmidt, Ann Marie
 APPLICANT: Lamster, Ira
 TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
 FILE REFERENCE: 0575/64080
 CURRENT APPLICATION NUMBER: US/09/872,185B
 CURRENT FILING DATE: 2001-06-01
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patent In version 3.1
 SEQ ID NO 5
 LENGTH: 30
 TYPE: PRT
 ORGANISM: Rat
 US-09-872-185B-5

Query Match: 22.4%, Score 133, DB 10; Length 30;
 Best Local Similarity: 82.8%, Pred. No. 2.8e-06;
 Matches: 24; Conservative: 2; Mismatches: 3; Indels: 0; Gaps: 0;

QY 2 QNITARIGEPVLKCKGAPKPKPPQRLWK 30
 DB 2 QNITARIGEPVLKCKGAPKPKPPQRLWK 30

Search completed: May 30, 2003, 15:59:06
 Job time: 9.82833 secs



Result No.	Query			ID	Description
	Score	Match	Length		
1	595	100	404	1	advanced glycosyl transferase
2	516.5	86	8	2	probable advanced glycosyl transferase
3	513.5	86	3	1	advanced glycosyl transferase
4	114	19.2	1416	1	tumor suppressor p
5	108	18.2	267	2	probable tumor suppressor
6	103	17.3	1147	2	myosin-light-chain
7	101	17.0	1176	2	myosin-light-chain
8	97	16.3	1040	2	transient axonal growth
9	96.5	16.2	1257	1	neural cell adhesion
10	96	16.1	1018	2	neural cell adhesion
11	96	16.1	1040	2	axonal glycoprotein
12	96	16.1	1344	2	rig 1 protein - mo
13	95.5	16.1	2295	2	hypothetical protein
14	95.5	16.1	3375	2	hypothetical protein
15	95	16.0	1823	2	pregnancy specific
16	95	16.0	1427	2	tumor suppressor
17	93	15.6	1018	2	contactin 1 precursor
18	92.5	15.6	1612	2	duffy protein - mo
19	92.5	15.5	1651	2	transmembrane receptor
20	92	15.5	732	2	transmembrane receptor
21	91	15.3	523	2	neurofilament
22	91	15.3	739	2	neurofilament
23	91	15.3	1535	2	vascular cell adhesion
24	89	15.0	1320	2	peroxidasin - fruit
25	88	14.8	332	2	neural cell surface
26	88	14.9	1033	2	pregnancy specific
27	87	14.6	210	2	cell adhesion protein
28	87	14.6	273	2	pregnancy specific
29	87	14.6	275	2	pregnancy specific

Best Local Similarity 100.0%; Pred. No. 66-52;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 60
DB 23 AONTARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 82

QY 61 SLFPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 112
DB 83 SLFPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 134

RESULT 2
T09062
Probable advanced glycosylation end-products receptor precursor - mouse
N:Alternate names: RAGE
C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun 1999 #text_change 21-Jan-2000
C:Accession: T09062
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahern, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; S.
submitted to the EMBL data library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: Z16543
A:Accession: T09062
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1 402 <ROW>
A:Cross-references: EMBL:AF030001; NID:q2564945; PID:q2564950
C:Genetics:
A:Gene: RAGE
A:Map position: 17
A:Introns: 18/2, 53/3, 118/3, 169/4, 209/4, 320/4, 329/4, 330/4, 371/2
A:Superfamily: advanced glycosylation end-products receptor; immunoglobulin homology
C:Keywords: receptor; transmembrane protein
F:31-100/Domain: immunoglobulin homology <IMM>

Query Match 86.8%; Score 516.5; DB 2; Length 402;
Best Local Similarity 88.3%; Pred. No. 43-44;
Matches 98; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 2 ONITARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 61
DB 24 ONITARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 82

QY 62 LFLPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 111
DB 83 LFLPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 142

RESULT 4
A54100
Tumor suppressor protein DCC precursor human
N:Alternate names: colorectal cancer suppressor DCC
C:Species: Homo sapiens (man)
C:Date: 02-Aug-1994 #sequence_revision 02-Aug 1994 #text_change 05-Nov-1999
C:Accession: A54100; A40098
R:Bedrick, L.; Cho, K.F.; Fearon, F.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
Genes Dev. 8, 1174-1183, 1994
A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis
A:Reference number: A54100; MUID:95011532; PMID:7926722
A:Accession: A54100
A:Molecule type: mRNA
A:Residues: 1-1447 <BED>
A:Cross-references: EMBL:X76132; NID:q454209; PID:CAA3745.1; PID:q454210
R:Fearon, F.R.; Cho, K.R.; Nigro, J.M.; Kern, S.F.; Simons, J.W.; Ruppert, J.M.; Hami
Science 247, 49-56, 1990
A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancer
A:Reference number: A40098; MUID:90100559; PMID:2294591
A:Accession: A40098
A:Molecule type: mRNA
A:Residues: 1-750 <FEA>
A:Cross-references: GB:M32292; NID:q181492; PID:AAA4575.1; PID:q181494
C:Genetics:
A:Gene: DCC
A:Cross-references: GDB:119838; OMIM:120470
A:Map position: 18q21.1-18q21.1
C:Keywords: transmembrane protein; tumor suppressor
F:1-25/Domain: signal sequence #status predicted -SIG-
F:26-1447/Product: tumor suppressor protein DCC #status predicted -MAT-

Query Match 19.2%; Score 114; DB 2; Length 1447;
Best Local Similarity 31.3%; Pred. No. 0-0047;
Matches 31; Conservative 20; Mismatches 48; Indels 10; Gaps 4;

QY 2 ONITARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 61
DB 147 ESVTAFMGDTVLKCE-VIGEPMTTHWQKNQ -QULTP1 -TCDSHAVVLPSSA 196

Best Local Similarity 100.0%; Pred. No. 66-52;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 60
DB 23 AONTARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 82

QY 61 SLFPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 112
DB 83 SLFPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 134

RESULT 2
T09062
Probable advanced glycosylation end-products receptor precursor - mouse
N:Alternate names: RAGE
C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun 1999 #text_change 21-Jan-2000
C:Accession: T09062
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahern, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; S.
submitted to the EMBL data library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: Z16543
A:Accession: T09062
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1 402 <ROW>
A:Cross-references: EMBL:AF030001; NID:q2564945; PID:q2564950
C:Genetics:
A:Gene: RAGE
A:Map position: 17
A:Introns: 18/2, 53/3, 118/3, 169/4, 209/4, 320/4, 329/4, 330/4, 371/2
A:Superfamily: advanced glycosylation end-products receptor; immunoglobulin homology
C:Keywords: receptor; transmembrane protein
F:31-100/Domain: immunoglobulin homology <IMM>

Query Match 86.8%; Score 516.5; DB 2; Length 402;
Best Local Similarity 88.3%; Pred. No. 43-44;
Matches 98; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 2 ONITARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 61
DB 24 ONITARIGEPVLVKCAKAPKPPQLEWKLNTGRTEAKVKLSPOGGGHWDSVARVLPG 82

QY 62 LFLPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 112
DB 83 LFLPAVGIDEGIFRCQAMNRNGKTSNYRVRVVOIPGKPEIVDSASELT 133

RESULT 4
A42879
advanced glycosylation end-products receptor precursor - bovine
N:Alternate names: advanced glycosylation end-product-binding protein, 35K; glycoprotein
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C:Accession: A42879; A42879; S27949
R:Neuper, M.; Schmidt, A.M.; Bretz, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; S.
J. Biol. Chem. 267, 14998-15004, 1992
A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation or
A:Reference number: A42879; MUID:92340547; PMID:1378843
A:Accession: A42879
A:Molecule type: mRNA
A:Residues: 1 416 <NEE>
A:Cross-references: GB:M91212; NID:q163650; PID:AAA03575.1; PID:q163651
A:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBI:109446)
R:Schmidt, A.M.; Vianna, M.; Gerlach, M.; Bretz, J.; Ryan, J.; Kao, J.; Esposito, C.; Be
J. Biol. Chem. 267, 14987-14997, 1992
A:Title: Isolation and characterization of two binding proteins for advanced glycosylat
A:Reference number: A42878; MUID:92340546; PMID:1321822
A:Accession: A42878
A:Molecule type: protein

QY 62 LFLFAVGQDEGIFPCQAMNNGKETKSNYKRVVYQIP 100
 DB 197 LQSPIDQDQICVRCAPNPFASRRICNCAEVEHSDPG 235

RESULT 5

A38442
 probable tumor suppressor - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 18-Feb-1997
 C:Accession: A38442
 R:Niaro, J.M.; Chu, K.P.; Fearon, E.R.; Karin, S.E.; Ruppert, J.M.; Gillen-r, J.D.; Kinzler
 Cell 64, 607-613, 1991
 A:Title: Scrambled exons.
 A:Reference number: A38442; MUID:91121517; PMID:1991322
 A:Accession: A38442
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <NT>
 A:Cross-references: GH:M63696; GH:M63700; GH:M63702; GH:M63718; GH:M63698
 C:Keywords: transmembrane protein

Query Match 18.2%, Score 108, DB 2, Length 267,
 Best Local Similarity 33.6%, Pred. No. 0.0022,
 Matches 30; Conservative 26; Mismatches 38; Indels 10, Gaps 3.

QY 2 QNITARTGEPVLVKCKGAPKKPPQRLWKLN--TGRTAANKVLSQGGGPDWSVARVLPN 61
 DB 40 ESVTAFMGDTVLKKE-VIGPEPTTHWKNQ-----QULTPI---PGDSRVVVLPSGA 89
 QY 62 LFLFAVGQDEGIFPCQAMNNGKETKSNYKRVVYQIP 99
 DB 90 LQSPIDQDQICVRCAPNPFASRRICNCAEVEHSDPG 127

RESULT 6

A59307
 myosin-light-chain kinase (BC 2.7.1.117), smooth muscle rabbit
 C:Species: Cytolagus olerulus (domestic rabbit)
 C:Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
 C:Accession: A59307; A41674; B41674; A40210
 R:Gallagher, P.J.; Herring, B.P.; Griffin, S.A.; Stull, J.T.
 J. Biol. Chem. 267, 23946-23948, 1991
 A:Title: Molecular characterization of a mammalian smooth muscle myosin light chain kinase
 A:Reference number: A41674; MUID:92084694; PMID:1748666
 A:Accession: A59307
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1147 <GA>
 A:Cross-references: GH:M76233, NID:j165703, PIDD:AAA73093 1, PIDD:j165704
 A:Experimental source: sex stage Adult; sex Female; tissue type smooth muscle
 A:Note: this revision submission is not cited in Genbank entry PARSMMLCKP, release 115
 A:Accession: A41674
 A:Molecule type: mRNA
 A:Residues: 1-738, 'RO', 741-1147 <GA>
 A:Cross-references: GH:M76233
 A:Note: the sequence is revised in Genbank entry PARSMMLCKP, release 115. (PIDD:AAA73093 1, PIDD:j165704)
 A:Accession: B41674
 A:Molecule type: DNA
 A:Residues: 1-22, 'CTA', 26, 'GFS', 'GAG'
 A:Cross-references: GH:M76369
 A:Note: this translation is not annotated in Genbank entry PARSMMLCKP, release 115
 A:Note: this sequence has been revised in reference A40210
 R:Gallagher, P.J.; Herring, B.P.; Griffin, S.A.; Stull, J.T.
 J. Biol. Chem. 267, 9450, 1992
 A:Reference number: A40210; MUID:92250555; PMID:1577772
 A:Contents: erratum
 A:Accession: A40210
 A:Molecule type: DNA
 A:Residues: 23-30 <GA>
 A:Note: this is a revision to the sequence B41674 from reference A41674
 C:Superfamily: unassigned Ser/Thr or Tyr specific protein kinases; fibronectin type III

C:Keywords: ATP, phosphotransferase, smooth muscle
 F:343-403/Domain: immunoglobulin homology <IMM>
 F:634-951/Domain: protein kinase homology <KIN>
 F:702-710/Region: protein kinase ATP-binding motif
 F:1055-1116/Domain: immunoglobulin homology <IMM2>

Query Match 17.3%, Score 103, DB 2, Length 1147,
 Best Local Similarity 27.3%, Pred. No. 0.036;
 Matches 30; Conservative 21; Mismatches 39; Indels 20; Gaps 5;

QY 2 QNITARTGEPVLVKCKGAPKKPPQRLWKLN--TGRTAANKVLSQGGGPDWSVARVLPN 59
 DB 346 QDVPVABSEKLLQGLQ-VSEPPATITWTNGKTLTKTFVLSQEGSLCSVSIEKALP- 393
 QY 60 GSLLFLPAVGQDEGIFPCQAMNNGKETKSNYKRVVYQIPGK-----PEI 104
 DB 394 -----EDRGLYKCVAKNA-APEAKSCHVTVDAPASENAKAPEM 432

RESULT 7

JN0583
 myosin-light-chain kinase (PI 2.7.1.117) jsek protein - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 03-Feb-1974 #sequence_revision 03-Feb-1994 #text_change 20-Apr 2000
 C:Accession: JN0583; PNU490
 R:Kikayashi, H.; Inoue, A.; Mikawa, T.; Kuwayama, H.; Hotta, Y.; Masaki, T.; Ebashi,
 J. Biochem. 112, 786-791, 1992
 A:Title: Isolation of cDNA for bovine stomach 155kDa protein exhibiting myosin light
 A:Reference number: JN0583; MUID:93203148; PMID:1284247
 A:Accession: JN0583
 A:Molecule type: mRNA
 A:Residues: 1-1176 <KOB>
 A:Cross-references: GR:S57331; NID:g298638; PIDD:AA825754.1; PIDD:g298639
 A:Experimental source: stomach
 A:Accession: PNU490
 A:Molecule type: protein
 A:Residues: 44-55, 721-728, 828-851, 1002-1019 <K02>
 C:Superfamily: unassigned Ser/Thr or Tyr specific protein kinases; fibronectin type I
 C:Keywords: ATP, calmodulin binding; phosphotransferase
 F:370-440/Domain: immunoglobulin homology <IMM>
 F:724-999/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:1083-1144/Domain: immunoglobulin homology <IMM2>

Query Match 17.0%, Score 101, DB 2, Length 1176;
 Best Local Similarity 29.1%, Pred. No. 0.058;
 Matches 30; Conservative 21; Mismatches 36; Indels 16; Gaps 5;

QY 2 QNITARTGEPVLVKCKGAPKKPPQRLWKLN--TGRTAANKVLSQGGGPDWSVARVLPN 59
 DB 363 QDLHVAE50KLLIQGLQ-VSSDPTAIIITWINGKILKTRKFTVLSQEGSLCSVSIEKALP- 420
 QY 60 GSLLFLPAVGQDEGIFPCQAMNNGKETKSNYKRVVYQIPGK 102
 DB 421 -----EDRGLYKCVAKNSAG-QAESSCQTV-DVPDAP 451

RESULT 8

A49356
 transient axonal glycoprotein TAG-1 precursor - human
 N:Alternate names: axonin-1
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 23-Mar-1995 #text_change 24-Sep-1999
 C:Accession: S35508; S28830; A49356
 R:Hasler, T.
 submitted to the EMBL Data Library, September 1992
 A:Reference number: S35508
 A:Accession: S35508
 A:Molecule type: mRNA
 A:Residues: 1-1040 <HAS>
 A:Cross-references: EMBL:X68274, NID:g36674, PIDD:CAA48335.1, PIDD:g36675
 R:Hasler, T.H.; Rader, C.; Stoeckli, E.T.; Zuehlke, P.A.; Sonderegger, P.
 Eur. J. Biochem. 211, 339-359, 1993

A:Title: cDNA cloning, structural features, and eucaryotic expression of human TAG-1/axd
A:Reference number: S28830; MUID:93145965; PMID:8425542
A:Accession: S28830
A:Molecule type: mRNA
A:Residues: 1-296, T, 298-1040 <HA>
A:Cross-references: EMBL:X68274
R:Tsotra, P.C.; Karagogeos, D.; Theodorakis, K.; Michaelidis, T.M.; Modi, W.S.; Furley, Genomics 18, 562-567, 1993
A:Title: Isolation of the cDNA and chromosomal localization of the gene (TAG1) encoding
A:Reference number: A49356; MUID:94140354; PMID:8307567
A:Accession: A49356
A:Molecule type: mRNA
A:Residues: 1-1001, G, 1003-1040 <HS1>
A:Cross-references: GH:X67734
C:Genetics:
A:Gene: GDB:TAX1
A:Cross-references: GDB:138782
A:Map position: 1q32-1q32
C:Superfamily: contactin; fibronectin type III repeat homology, immunoglobulin homology
C:Keywords: cell adhesion; glycoprotein; predicted <SIG>
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1040/Product: axonal glycoprotein TAG 1 #status predicted <MAT>
F:254-308/Domain: immunoglobulin homology <IMM1>
F:341-397/Domain: immunoglobulin homology <IMM2>
F:76, 198, 204, 461, 477, 498, 525, 775, 830, 904, 918, 940/Binding site: carbohydrate (covar)
Query Match 16.4%; Score 97; DB 2; Length 1040;
Best local similarity 40.5%; Pred. No. 0.13;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;
QY 6 ARICEPLVVKCKAPKPPQRLKWLKNTGTEAWKVI SPQGGGHWDSVARVPLNGSLFLP 65
DB 251 ALVGGQVTEC-FAFGNVPRIKRWKVDGS-----LSPQ---WTTA-----EPTLQIP 294
QY 66 AVGLQDEGIFRCQANNRCKET 87
DB 295 SVSFEDEGTVECAFNKSGRDT 316
RESULT 9
neural cell adhesion molecule L1 precursor - human
A:1060
N:Alternate names: L1CAM
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence revision 10 Sep 1999 #text change 31 Jul 2000
C:Accession: A41060; S18454; A35331; S21971; S21972; A60223; A31072; G02506
R:Blavin, M.L.; Lemmon, V.
A:Title: Molecular structure and functional testing of human L1CAM: an interspecies comp
A:Reference number: A41060; MUID:92120663; PMID:1769655
A:Accession: A41060
A:Molecule type: mRNA
A:Residues: 1-1257 <L1>
A:Cross-references: GH:M64296; NID:q186053; PIDN:AA014352.1; PID:q3068548
R:Kobayashi, M.; Miura, M.; Asou, H.; Uyemura, K.
Biochim. Biophys. Acta 1090, 238-240, 1991
A:Title: Molecular cloning of cell adhesion molecule L1 from human nervous tissue: a com
A:Reference number: S18454; MUID:92031698; PMID:1942117
A:Accession: S18454
A:Molecule type: mRNA
A:Residues: 1-3, V, 5, 215, 1, 217-249, 1, 251 275, 3V, 278 356, 4V, 458 625, 6V, 627-1257
A:Cross-references: EMBL:X59847; NID:q45009; PIDN:CAA42508.1; PID:q35010
A:Note: the authors translated the codon GAA for residue 27 as Gly
R:Diabali, M.; Maitel, M.G.; Nguyen, C.; Roux, D.; Demenqot, J.; Denizot, F.; Moos, M.; Genomics 7, 587-593, 1990
A:Title: The gene encoding L1, a neural adhesion molecule of the immunoglobulin family,
A:Reference number: A35331; MUID:90353957; PMID:2387585
A:Accession: A35331
A:Molecule type: DNA
A:Residues: 332-371 <DJA>
A:Cross-references: GH:M55271
R:Rosenthal, A.; MacKinnon, R.N.; Jones, D.S.C.
Nucleic Acids Res. 19, 5395-5401, 1991
A:Title: PCR walking from microdissection clone M54 identifies three exons from the b
A:Reference number: S21971; MUID:92020233; PMID:1923824
A:Accession: S21971
A:Molecule type: DNA
A:Residues: 1082-1176 <ROS>
A:Cross-references: EMBL:X58775; NID:q29642; PIDN:CAA41376.1; PID:q296443
A:Accession: S21972
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 353-935, V, 937-1176 <R02>
A:Cross-references: EMBL:X58776; NID:q29644; PIDN:CAA43781.1; PID:q4467844
R:Harper, J.R.; Prince, J.T.; Healy, P.A.; Stuart, J.K.; Nauman, S.J.; Stallcup, W.R.
J. Neurochem. 56, 797-804, 1991
A:Title: Isolation and sequence of partial cDNA clones of human L1: homology of human
A:Reference number: A60223; MUID:91132183; PMID:1993895
A:Accession: A60223
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1030-1115, W, 1116-1176, 1181-1257 <HAR>
R:Wolff, J.M.; Frank, R.; Mujoo, K.; Spiro, R.C.; Reistfeld, R.A.; Nathjien, F.G.
J. Biol. Chem. 263, 11943-11947, 1988
A:Title: A human brain glycoprotein related to the mouse cell adhesion molecule L1.
A:Reference number: A31072; MUID:88298876; PMID:3146168
A:Accession: A31072
A:Molecule type: protein
A:Residues: 1-1257 <WOL>
R:Platzter, M.; Bauer, D.; Drescher, B.
submitted to the EMBL Data Library, March 1995
A:Reference number: H01368
A:Accession: G02506
A>Status: preliminary; translated from GH/EMBL/DBR3
A:Molecule type: DNA
A:Residues: 1-1257 <PLA>
A:Cross-references: EMBL:052112; NID:q1302657; PIDN:AA01746.1; PID:q1402658
C:Genetics:
A:Gene: GDB:L1CAM
A:Cross-references: GDB:120133; OMIM:403450; OMIM:408840
A:Map position: Xq28-Xq28
A:Introns: 26/1, 31/1, 56/2, 134/1, 175/1, 196/2, 33/3, 37/2, 43/1, 45/2;
/2
C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology
C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; transmem
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1257/Product: neural cell adhesion molecule L1 #status predicted <MAT>
F:257-314/Domain: immunoglobulin homology <IMM1>
F:542 599/Domain: immunoglobulin homology <IMM2>
Query Match 16.2%; Score 96.5; DB 1; Length 1257;
Best local similarity 32.4%; Pred. No. 0.18;
Matches 45; Conservative 12; Mismatches 40; Indels 21; Gaps 4;
QY 1 AONTATATGEGPLVLRK--KGAPKPPQRLKWLKNTGTEAWKVI SPQGGGHWDSVARVPL 58
DB 249 SSSHVALQGPVLECLANGP-----GPPFLKLRPSGPMIAHRTVYQNH 293
QY 59 NCSFLPAPVQIDEGIFRCQANNRCKET KSNRYRVVYQID GKPE 104
DB 294 NKTQLLKVSEEDGEYPTAFNSLG-SARIAYVTVTEAAYVLLKRPQ 440
RESULT 10
neural adhesion protein F3 precursor bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Nov-1995 #sequence revision 08-Feb 1996 #text change 20 Jun 2000
C:Accession: JC4211
R:Watanabe, K.; Shimazaki, K.; Hosoya, H.; Fukumachi, F.; Ikenawa, T.
Gene 160, 245-248, 1995
A:Title: Cloning of the cDNA encoding neural adhesion molecule F3 from bovine brain.
A:Reference number: JC4211; MUID:95369697; PMID:7642105
A:Accession: JC4211
A:Molecule type: mRNA
A:Residues: 1 1018 <MAT>

A:Cross-references: UniProtKB:P51553; Nidoglobulin; P10N-AAA-00001; P10N000061
 A:Experimental source: brain
 C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
 C:Keywords: brain; contactin; glycoprotein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-1018/Region: hydrophobic #status predicted <IMM>
 F:226-312/Domain: immunoglobulin homology
 F:992-1018/Region: hydrophobic #status predicted
 F:208-258,348,457,473,494,521-591,630,933/Binding site: carbohydrate (Asn) (covalent) #

Query Match 16.18; Score 96; DB 2; Length 1018;
 Best local Similarity 26.48; Pred. No. 0.16;
 Matches 29; Conservative 26; Mismatches 35; Indels 20; Gaps 5;

QY 2 QNITAPIGERVLVKQKAFKPKPQLEWKINTGTEAWKVLSPQGGPQWDSVAR 61
 DB 249 KDVALMGQNTLECFALGNPFDIRWR-----KVLEPM-----PSTAEISTGA 293
 QY 62 LPTPAVGQDGFRCQAMNNGKETKSNYRVPVYCIQKPIVNSASE 110
 DB 294 VLKTFNQLDEGGIYECEAENNRKQD---KIKAKIV-VQAFPEVWEHIND 339

RESULT 11
 A34695
 axonal glycoprotein TAG-1 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 29-Jun-1970 #sequence_revision 29-Jun-1990 #text_change 21-Jan-2000
 C:Accession: A34695
 R:Furley, A.J., Morton, S.B., Manalo, D., Karagogeos, D., Dodd, J., Jessell, T.M.
 Cell 61, 167-178, 1990
 A:Title: The axonal glycoprotein TAG-1 is an immunoglobulin superfamily member with novel
 A:Reference number: A34695; MIM:601988N0; PMID:2417872
 A:Accession: A34695
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1848 <R>
 A:Cross-references: GB:M1725; UniProt:P207148; P10N-AAA12001; P10N0207149
 C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
 C:Keywords: glycoprotein
 F:343-399/Domain: immunoglobulin homology <IMM>

Query Match 16.18; Score 96; DB 2; Length 1040;
 Best local Similarity 30.58; Pred. No. 0.16;
 Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 APTEFFELVVKKAFKPKPQLEWKINTGTEAWKVLSPQGGPQWDSVAR 55
 DB 253 ALVGGQVTLQCFALGNPFDIRWR-----LSQD---WATA-----EPHQLIP 296
 QY 66 AVGLQDEGIFRCQAMNNGKET 87
 DB 297 SVSFEDEGTVECEAENSKGRDT 318

RESULT 12
 T14316
 T14316 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Sep-1994 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14316
 R:Yuan, S.F.; Cox, I.A.; Basik, A.; Lee, E.Y.H.P.
 submitted to the EMBL Data Library, April 1998
 A:Reference number: T17975
 A:Accession: T14316
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-1344 <Y>
 A:Cross-references: EMBL:AF060570; NID:14206385; P10N-AAA11628 1

Query Match 16.18; Score 96; DB 2; Length 1344;
 Best local Similarity 27.78; Pred. No. 0.21;
 Matches 31; Conservative 11; Mismatches 34; Indels 30; Gaps 5;

QY 9 GEPVLKCKGAPKPKPQLEWKINTGTEAWKVLSPQGGPQWDSVAR-----VLPNG 60
 DB 56 GEPATLPCR-AEGPPPPNIEWYKNGAPV-----ATAPEDPPAHPILLPSG 99
 QY 61 SIPDAV-----GLVDEGIFRCQAMNNGKETKSNYKVKV-----YQLPG 100
 DB 100 ALFFPR-VHGRPSKPGHGVYVYVAKNYTAAASPNASLEAVVLPIDFROSPG 151

RESULT 13
 C88369
 C88369: unc-52 (unpooled - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 18-Jul-2001
 C:Accession: C88369; E88369
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequences of the nematode C. elegans: a platform for investigating bio
 A:Reference number: A75000; MIM:590596I3; PMID:9851916
 A:Note: see websites genome.wustl.edu/qsc/C_elegans/ and www.sanger.ac.uk/Projects/C-
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A:Accession: C88369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2295 <S>
 A:Cross-references: GB:chr_11; P10N-CAB07568.1; P10N-9874868; GSPUB:GN000020
 A:Accession: C88369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2295 <S>
 A:Cross-references: GB:chr_11; P10N-CAB07568.1; P10N-9874868; GSPUB:GN000020
 C:Genetics:
 A:Gene: unc-52
 A:Map position: 2

Query Match 16.18; Score 95.5; DB 2; Length 2295;
 Best local Similarity 25.48; Pred. No. 0.44;
 Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;

QY 8 IGEPVLKCKK-GAPKKPKPQLEWKINTGTEAWKVLSPQGGPQWDSVAR 54
 DB 1803 VGHPIQKCKAFKPKPQLEWKINTGTEAWKVLSPQGGPQWDSVAR 1853
 QY 65 PAVGLQDEGIFRCQAMNNGKETKSNY-----RVRYVQIRG-----KPEIVDS 107
 DB 1854 PNVGLSNAGVYICKGSAHATAIKNIYIEWVEPKRIATVILGGSSQWELQGEKGLICT 1913
 QY 108 AS 109
 DB 1914 AT 1915

RESULT 14
 T19821
 hypothetical protein ZC101.2e - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
 C:Accession: T19821; T19820; T27490; T27488; T27489; T27487; A47648; B47648;
 R:Baynes, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19182
 A:Accession: T19821
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-3375 <W>
 A:Cross-references: EMBL:Z43375; P10N-CAB07569.1; GSPUB:GN000020; CESP:ZC101.2e
 A:Experimental source: clone C38C6
 A:Accession: T19819
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-2441 <K>; KKKH:33369; GN:3372-3373; G:3375; RLRHRRNAONGPLSKRTTTTKLF
 A:Cross-references: EMBL:Z43375; P10N-CAB07567.1; GSPUB:GN000020; CESP:ZC101.2a

GenCore version 5.1
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:37:16, Search time 6.55856 seconds
(without alignments)
708.288 million cell updates/sec

Title: US-09-872-185B-1

Perfect score: 595
Sequence: 1 AQNITARIGEPLVLCKGNP VPVYQIPKPEIVDSASELT 112

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	595	100.0	404	1	RACE_HUMAN	Q15109 homo sapien
2	513.5	86.3	416	1	RACE_BOVIN	Q28173 bos taurus
3	512.5	86.1	403	1	RACE_MOUSE	Q62151 mus musculus
4	494.5	83.1	402	1	RACE_RAT	Q63495 rattus norv
5	115	19.3	1447	1	DCC_MOUSE	P70211 mus musculus
6	114	19.2	1447	1	DCC_HUMAN	P43146 homo sapien
7	103	17.3	1176	1	KMIS_PARIT	P29234 oryctolagus
8	101	17.0	1176	1	KMIS_BOVIN	Q28824 bos taurus
9	97	16.3	1040	1	AXOL_HUMAN	Q02246 homo sapien
10	96.5	15.2	1257	1	CAML_HUMAN	P32004 homo sapien
11	96	16.1	1040	1	AXOL_RAT	P22063 rattus norv
12	95.5	15.7	2481	1	UN52_CAEFL	Q06501 caenorhabdi
13	93.5	15.7	713	1	GACL_HUMAN	O75325 homo sapien
14	93	15.6	1018	1	CONT_HUMAN	Q12860 homo sapien
15	92	15.5	702	1	CEAS_HUMAN	Q06731 homo sapien
16	91	15.3	555	1	CL66_CARAU	Q90304 carassius a
17	91	15.3	739	1	VCAI_RAT	P29534 rattus norv
18	90.5	15.2	1914	1	KMIS_HUMAN	Q15746 homo sapien
19	89	15.0	1020	1	CONT_MOUSE	P12960 mus musculus
20	87	14.6	3707	1	PGBM_MOUSE	Q05793 mus musculus
21	86.5	14.5	1266	1	NGCA_CHICK	Q03696 gallus gall
22	86	14.5	1021	1	CONT_RAT	Q63198 rattus norv
23	85.5	14.4	4393	1	PGBM_HUMAN	P98160 homo sapien
24	85	14.3	1010	1	CONT_CHICK	P14781 gallus gall
25	85	14.3	1036	1	AXOL_CHICK	P28685 gallus gall
26	84.5	14.2	337	1	G55A_CHICK	Q98892 gallus gall
27	84	14.1	873	1	FAS2_DROME	P34082 drosophila
28	84	14.1	1356	1	VGR2_HUMAN	P35968 homo sapien
29	82.5	13.9	338	1	LAMP_RAT	Q62813 rattus norv
30	82.5	13.9	345	1	OPCM_HUMAN	Q14982 homo sapien
31	82.5	13.9	345	1	OPCM_RAT	P32736 rattus norv
32	82	13.8	673	1	SM4B_HUMAN	Q9NPR2 homo sapien
33	82	13.8	850	1	NRG2_HUMAN	Q14511 homo sapien

RESULT 1

RAGE_HUMAN	34	41.5	13.7	345	1	OPCM_BOVIN	P11834 bos taurus
ID	35	41.5	13.7	782	1	SM4B_MOUSE	Q62179 mus musculus
AC	36	81	13.6	898	1	FAS2_SCHAM	P26448 schistocerc
DT	37	80.5	13.5	338	1	LAMP_HUMAN	Q13449 homo sapien
DT	38	80.5	13.5	1336	1	VGR1_RAT	P53767 rattus norv
DT	39	79	13.3	419	1	PSG1_HUMAN	P11464 homo sapien
DE	40	78.5	13.2	338	1	LAMP_CHICK	Q98919 gallus gall
DE	41	78.5	13.2	756	1	NRG2_MOUSE	P56974 mus musculus
DE	42	78.5	13.2	868	1	NRG2_RAT	O35569 rattus norv
DE	43	78.5	13.2	999	1	MERK_HUMAN	O12866 homo sapien
GN	44	78.5	13.2	1333	1	VGR1_MOUSE	P35969 mus musculus
OS	45	78	13.1	994	1	MERK_RAT	P57097 rattus norv

STANDARD: PPT: 404 AA.

Q15109; Q15279; Q9Y3R3; Q9H2X7;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Advanced glycosylation end product-specific receptor precursor
(Receptor for advanced glycosylation end products)
AGER OR RAGE.
Homo sapiens (Human).
Fukuyota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Lung;
MEDLINE=92340547; PubMed=1378843;
Neeper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
Elliston K., Stern D., Shaw A.;
Cloning and expression of a cell surface receptor for advanced
glycosylation end products of proteins.*;
J. Biol. Chem. 267:14998-15004(1992).
[2]
SEQUENCE FROM N.A. (ISOFORM 1).
MEDLINE=95137587; PubMed=7835890;
Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,
Inoko H., Ikemura T.;
Three genes in the human MHC class III region near the junction with
the class II: gene for receptor of advanced glycosylation end
products, PBX2 homeobox gene and a notch homolog, human counterpart
of mouse mammary tumor gene int-3.*;
Genomics 23:408-419(1994).
[3]
SEQUENCE FROM N.A. (ISOFORM 1).
Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
Banta A., Spies T., Hood L.;
Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A. (ISOFORM 1). AND VARIANT APC-100.
Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,
Yamamoto H.;
Molecular heterogeneity of the receptor for advanced glycation
endproducts.*;
Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A. (ISOFORM 2).
Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,
Schuler A., Huber G.;
cDNA cloning of a novel secreted isoform of the human Receptor for
Advanced Glycation End products (RAGE) and characterization of cells
co-expressing cell-surface scavenger receptors and Swedish mutant
amyloid precursor protein.*;
Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A. (ISOFORM 1).

ALIGNMENTS

FT DOMAIN 136 214 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 262 318 IG-LIKE C2-TYPE DOMAIN 2.
 FT DISULFID 38 98 POTENTIAL.
 FT DISULFID 143 207 POTENTIAL.
 FT DISULFID 269 311 POTENTIAL.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DOMAIN 391 396 POLY-GLU.
 SQ SEQUENCE 416 AA; 44182 MW; 870381557347677AR CQC64;

 Query Match 86 18; Score 513.5; DB 1; Length 416;
 Best Local Similarity 89 18; Pred. No. 2.4e-46;
 Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

 QY 2 QNITARIGELVLKCKGAPKPPORLEWKLNTGRTRFAWKVLSPOGGPWDSVARVLPNGS 61
 DB 24 QNITARICKPLVLNCKGAPKPPQQLLEWKLNTGRTRFAWKVLSPO-GDPWDSVARVLPNGS 82

 QY 62 LFLPAVILDEHIFFCVAMNPNKSKTSNYRVVYQIFGKPEIVDSASEL 111
 DB 83 LLLPAVGIGVDEGTFECATNRKCKEVSRYRVVYQIFGKPEIVDPASEL 132

 RESULT 3
 RAGE_MOUSE
 ID PAGE_MOUSE STANDARD: PRT: 403 AA.
 AC Q62151;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Advanced glycosylation end product-specific receptor precursor
 DE (Receptor for advanced glycosylation end products).
 GN AGPR OR RAGE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SOURCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=Lung;
 RX MEDLINE=97368045; PubMed=9224812;
 PA Penard C., Chappoy O., Wautier J.M., Nagashima M., Lundh E.,
 PA Morser J., Zhao L., Schmidt A.M., Schermann J.M., Wautier J.L.,
 RT "Recombinant advanced glycation end product receptor pharmacokinetics
 RT in normal and diabetic rats".
 RL Mol. Pharmacol. 52:54-62(1997).
 CC -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 CC RATE IN DIABETES.
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L33412; AAA0040.1;
 DR MGD; MG1:893592; Ager.
 DR InterPro; IPR003006; Iq_MHC.
 DR InterPro; IPR003598; Iq_C2.
 DR InterPro; IPR003600; Iq_Like.
 DR Pfam; PF00047; Iq; 3.
 DR SMART; SM00410; Iq_Like; 1.
 DR SMART; SM00408; Iq_C2; 1.
 DR PROSITE; PS00290; Iq_MHC; 1.

KW Immunoglobulin domain, Glycoprotein, Transmembrane, Repeat; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 403 ADVANCED GLYCOSYLATION END PRODUCT-SPECIFIC RECEPTOR.
 FT DOMAIN 23 341 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 342 362 POTENTIAL.
 FT DOMAIN 363 403 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 31 105 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 136 214 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 251 307 IG-LIKE C2-TYPE DOMAIN 2.
 FT DISULFID 38 98 POTENTIAL.
 FT DISULFID 143 207 POTENTIAL.
 FT DISULFID 258 300 POTENTIAL.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 433 AA; 42668 MW; 1279796FD1579357 CRC64;

 Query Match 86 18; Score 512.5; DB 1; Length 403;
 Best Local Similarity 87 18; Pred. No. 2.9e-46;
 Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

 QY 2 QNITARIGELVLKCKGAPKPPORLEWKLNTGRTRFAWKVLSPOGGPWDSVARVLPNGS 61
 DB 24 QNITARIGELVLKCKGAPKPPQQLLEWKLNTGRTRFAWKVLSQ-GGPWDSVARVLPNGS 82

 QY 62 LFLPAVILDEHIFFCVAMNPNKSKTSNYRVVYQIFGKPEIVDSASEL 112
 DB 83 LLLPAVGIGVDEGTFECATNRKCKEVSRYRVVYQIFGKPEIVDPASEL 133

 RESULT 4
 RAGE_RAT
 ID RAGE_RAT STANDARD: PRT: 402 AA.
 AC Q63495;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Advanced glycosylation end product-specific receptor precursor
 DE (Receptor for advanced glycosylation end products).
 GN AGER OR RAGE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Lung;
 RX MEDLINE=97368045; PubMed=9224812;
 PA Penard C., Chappoy O., Wautier J.M., Nagashima M., Lundh E.,
 PA Morser J., Zhao L., Schmidt A.M., Schermann J.M., Wautier J.L.,
 RT "Recombinant advanced glycation end product receptor pharmacokinetics
 RT in normal and diabetic rats".
 RL Mol. Pharmacol. 52:54-62(1997).
 CC -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
 CC RATE IN DIABETES.
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L33413; AAA42027.1;
 DR InterPro; IPR003006; Iq_MHC.

Fri May 30 17:16:33 2003

```

DR InterPro: IPR003598; Iq_c2.
DR InterPro: IPR003600; Iq_1like.
DR Pfam: PF00047; Iq; 3.
DR SMART: SM00410; Iq_1like; 1.
DR SMART: SM00408; Iq_2; 1.
DR PROSITE: PS00290; Iq_MHC; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 402
FT DOMAIN 23 441
FT TRANSMEM 342 402
FT DOMAIN 363 402
FT DOMAIN 31 105
FT DOMAIN 136 213
FT DOMAIN 250 306
FT DOMAIN 48 98
FT DISULFID 143 206
FT DISULFID 257 299
FT CARBOHYD 25 25
FT CARBOHYD 80 80
SQ SEQUENCE 402 AA; 42663 MW; 594481BCA51E94FC CRC64;

Query Match 81.1%; Score 494.5; DB 1; Length 402;
Best Local Similarity 84.7%; Pred. No. 2 2a-44;
Matches 94; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 2 QNITARIGEPVLEKFKAPKPPQPLEWKINTGRTFAWKVLSPQSGPWDSVARVLPGNS 61
DB 24 QNITARIGEPVLEKFKAPKPPQPLEWKINTGRTFAWKVLSPQSGPWDSVARVLPGNS 82
QY 62 LFLPVGCIQDEGFPCQAMNPNKFKSKNYRVVYQIPGKPEIWDASSET 112
DB 83 LLLPALGIVDEGTFRCRATNRLGKEVKSNYRVVYQIPGKPEIWDASSET 133

RESULT 5
ID DCN_MOUSE
AC P70211;
DI 01-NOV-1997 (Rel. 35, created)
DI 15-JUN-2002 (Rel. 41, last annotation update)
DE tumor suppressor protein DCC precursor.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC STRAIN-BALB/C; TISSUE-Brain;
RX MEDLINE=96112625; PubMed=94576174;
RA Cooper H.M., Ames P., Britto J., Gad J., Wilks A.P.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RT gene (mDCC) and its expression in the developing mouse embryo.";
RL Genbank 11-2243-2254 (1995).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C; TISSUE-Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
CC produced by alternative splicing.
CC 1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
CC 1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION

```

```

CC AND REMAIN AT THIS LEVEL IN THE ADULT.
CC 1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC 1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS
CC 1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/submit/announce/
CC or send an email to license@ebi.ac.uk.)
CC
CC EMBL: X85788; CAA59786.1;
CC HSSP: P56276; ITLK.
CC MGD: MGI:94869; DCC.
DR InterPro: IPR004961; FN_111.
DR InterPro: IPR004962; FN111_repeat.
DR InterPro: IPR004006; Iq_MHC.
DR InterPro: IPR004598; Iq_1.
DR InterPro: IPR004600; Iq_1like.
DR Pfam: PF00041; Iq; 3.
DR Pfam: PF00047; Iq; 4.
DR PRINTS: PR00014; FN111.
DR SMART: SM00060; FN3; 6.
DR SMART: SM00410; Iq_1like; 2.
DR SMART: SM00408; Iq_2; 3.
DR SMART: SM00408; Iq_2; 3.
KW Glycoprotein; Immunoglobulin; domain; Transmembrane; Signal.
KW Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1447
FT CHAIN 85 1447
FT INIT MET 85 85
FT DOMAIN 26 1097
FT TRANSMEM 1098 1122
FT DOMAIN 1123 1447
FT DOMAIN 54 124
FT DOMAIN 154 219
FT DOMAIN 254 317
FT DOMAIN 345 407
FT DOMAIN 426 522
FT DOMAIN 525 618
FT DOMAIN 619 716
FT DOMAIN 722 816
FT DOMAIN 840 940
FT DOMAIN 941 1042
FT DISULFID 61 117
FT DISULFID 161 212
FT DISULFID 261 310
FT DISULFID 352 400
FT CARBOHYD 94 94
FT CARBOHYD 299 299
FT CARBOHYD 318 318
FT CARBOHYD 478 478
FT CARBOHYD 628 628
FT CARBOHYD 732 732
FT CARBOHYD 819 838
FT VARSPLIC 1447 AA; 158298 MW; 0D1F097C2D5H9F CRC64;
SQ SEQUENCE 1447 AA; 158298 MW; 0D1F097C2D5H9F CRC64;

Query Match 91.3%; Score 115; DB 1; Length 1447;
Best Local Similarity 91.3%; Pred. No. 0 00046;
Matches 31; Conservative 20; Mismatches 48; Indels 10; Gaps 3;

QY 2 QNITARIGEPVLEKFKAPKPPQPLEWKINTGRTFAWKVLSPQSGPWDSVARVLPGNS 61
DB 147 ESTAFMCDIVLKKE-VIGEPMTTHWQNG 196
QY 62 LFLPVGCIQDEGFPCQAMNPNKFKSKNYRVVYQIPGKPEIWDASSET 100
DB 197 LQISRIQPGDSGVYRCRSARNIVASIRTONIAEVRILSDPG 235

```

RESULT 6
 DCC_HUMAN STANDARD: PRT: 1447 AA.
 AC P43146;
 DT 01-NOV-1995 (rel. 32, corrected)
 DT 01-NOV-1995 (rel. 32, last sequence update)
 DT 15-JUN-2002 (rel. 41, last annotation update)
 DE Tumor suppressor protein DCC precursor (Colorectal cancer suppressor)
 GN DCC.
 OS Homo sapiens (Human).
 OC Eukaryota, Metazoa, Chordata, Cnidaria, Vertebrata, Euteleostomi;
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
 OX NCBI_taxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95011532; PubMed-2294591;
 RA Hedrick L., Cho K.R., Fearon F.R., Wu T.-C., Kinzler K.W.,
 RA Vogelstein B.;
 RA "The DCC gene product in cellular differentiation and colorectal
 RT tumorigenesis";
 RL Genes Dev. 8:1174-1183(1994)
 RN [2]
 RP SEQUENCE OF 1-750 FROM N.A.
 RX MEDLINE-90100559; PubMed-2294591;
 RA Fearon F.R., Cho K.R., Nigro J.M., Keln S.E., Simons J.W.,
 RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
 RA Vogelstein B.;
 RA "Identification of a chromosome 18q gene that is altered in
 RT colorectal cancers";
 RL Science 247:49-56(1990).
 RN [3]
 RP SEQUENCE OF 107-472 FROM N.A. (SCFAMBEID EXONS)
 RX MEDLINE-91121517; PubMed-1901322;
 RA Nigro J.M., Cho K.R., Fearon F.R., Keln S.E., Pappert J.M.,
 RA Oliner J.N., Kinzler K.W., Vogelstein B.;
 RA "Scrambled exons";
 RL Cell 64:607-613(1991)
 RN [4]
 RP GENE STRUCTURE AND VARIANTS (CARCINOMA HFS-1375)
 RX MEDLINE-94245241; PubMed-8188295;
 RA Cho K.R., Oliner J.N., Simons J.W., Hedrick L., Fearon F.R.,
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
 RA "The DCC gene: structural analysis and mutations in colorectal
 RT carcinomas";
 RL Genomics 19:525-531(1994).
 RN [5]
 RP VARIANT (CARCINOMA THR-158, AND VARIANT GLY-201).
 RX MEDLINE-94243823; PubMed-8187090;
 RA Miyake S., Nagai K., Yoshino K., Ota M., Endo M., Yusa Y.;
 RA "Point mutations and allelic deletion of tumor suppressor gene DCC in
 RT human esophageal squamous cell carcinomas and their relation to
 RT metastasis";
 RL Cancer Res. 54:3007-3010(1994)
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: type II membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
 CC DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMITY LACK DCC
 CC EXPRESSION. INACTIVATION OF DCC LEAD TO ALLELIC DELETION AND/OR
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS
 CC METASTASIS OF GASTROINTESTINAL SQUAMOUS CELL CARCINOMAS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY DCC
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X76132, GAA53735.1;
 CC EMBL: M32292, AAS35751.1;
 CC EMBL: M32286, AAS3174.1;
 CC EMBL: M32286, AAS3175.1; ALT_SEQ.
 CC EMBL: M32296, AAS2176.1;
 CC EMBL: M32296, AAS2177.1;
 CC EMBL: M63700, AAS2178.1;
 CC EMBL: M63702, AAS2179.1;
 CC EMBL: M63718, AAS2180.1;
 CC EMBL: M63698, AAS2181.1;
 CC EMBL: A54100, A54100;
 CC PIR: A40098; A40098;
 CC PIR: A38442; A38442;
 CC HSP: P56276; IILK;
 CC Genew: HGC:1701; DCC.
 CC MIM: 120470;
 CC InterPro: IP0003961; FN_III.
 CC InterPro: IP0003962; FNIII_repeat.
 CC InterPro: IP0003966; Iq_MHC.
 CC InterPro: IP0003968; Iq_C2.
 CC InterPro: IP0003600; Iq_Like.
 CC Pfam: PF00041; f63; 6.
 CC Pfam: PF00047; Iq; 4.
 CC PRINTS: PR00014; FNYPEIII.
 CC SMART: SM00060; FN3; 6.
 CC SMART: SM00410; Iq_Like; 2.
 CC SMART: SM00408; IqC2; 3.
 CC Glycoprotein; Immunoglobulin domain; Transmembrane; Signal,
 CC Repeat; Anti-oncogene; Disease mutation; Polymorphism.
 CC SIGNAL 1 25
 CC CHAIN 26 1447
 CC DOMAIN 26 1097
 CC TRANSMEM 1098 1122
 CC DOMAIN 1123 1447
 CC DOMAIN 124 124
 CC DOMAIN 154 219
 CC DOMAIN 254 317
 CC DOMAIN 345 407
 CC DOMAIN 426 522
 CC DOMAIN 525 618
 CC DOMAIN 619 716
 CC DOMAIN 722 816
 CC DOMAIN 840 940
 CC DOMAIN 941 1042
 CC DISULFID 51 117
 CC DISULFID 161 212
 CC DISULFID 251 310
 CC DISULFID 352 400
 CC CARBOHYD 94 94
 CC CARBOHYD 299 299
 CC CARBOHYD 318 318
 CC CARBOHYD 478 478
 CC CARBOHYD 628 628
 CC CARBOHYD 702 702
 CC VARIANT 168 168
 CC VARIANT 201 201
 CC VARIANT 1375 1375
 CC VARIANT 138 138
 CC CONFLICT 233 329
 CC CONFLICT 421 421
 CC CONFLICT 1447 1447
 CC SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;
 CC
 CC Query Match 19.2%; Score 114; DB 1; Length 1447;
 CC Best local similarity 31.3%; Pred. No. 0.00059;
 CC Matches 31; Conservative 20; Mismatches 38; Indels 10; Gaps 3;
 CC
 CC 2 QNTTARICEPLVTKKGAFFKPPQKLEWKLNTGRTFAWKVLSHQGGPWDSVARVLPGNS 61

[9] SEQUENCE OF 1030-1257 FROM N.A.
 RX MEDLINE-91132183; PubMed-1993895,
 RA Harper J.P., Prince J.T., Healy P.A., Nauman S.J.,
 RA Stallcup W.B.;
 RT "Isolation and sequence of partial cDNA clones of human L1: homology
 RT of human and rodent L1 in the cytoplasmic region.";
 RL J. Neurochem. 56:797-804(1991).
 [10]
 RP SEQUENCE OF 20-26.
 RX MEDLINE-88298876; PubMed-3136168;
 RA Wolff J.M., Frank R., Mujoo K., Spiro K.C., Weisfeld R.A.,
 RA Kachgen F.G.;
 RT "A human brain glycoprotein related to the mouse cell adhesion
 RT molecule L1.";
 RL J. Biol. Chem. 263:11943-11947(1988).
 [11]
 RP VARIANT HSAS TYR-264.
 RX MEDLINE-94004956; PubMed-8401576;
 RA Jouet M., Rosenthal A., Macfarlane J., Kenwright S., Donnai D.;
 RT "A missense mutation confirms the L1 defect in X-linked hydrocephalus
 RT (HSAS).";
 RL Nat. Genet. 4:331-331(1993).
 [12]
 RP VARIANT HSAS/MASA LEU-1194.
 RX MEDLINE-95187172; PubMed-7881431;
 RA Fransen E., Schrander-Stumpel C., Vits L., Coucke P., van Camp G.,
 RA Willems P.J.;
 RT "X-linked hydrocephalus and MASA syndrome present in one family are
 RT due to a single missense mutation in exon 28 of the L1CAM gene.";
 RL Hum. Mol. Genet. 3:2255-2256(1994).
 [13]
 RP VARIANTS HSAS GLN-184 AND ARG-452, AND VARIANT MASA GLN-210.
 RX MEDLINE-95004608; PubMed-7920659;
 RA Jouet M., Rosenthal A., Armstrong G., Macfarlane J., Stevenson R.,
 RA Paterson J., Metzberg A., Ionescu V., Temple K., Kenwright S.;
 RT "X-linked spastic paraplegia (SPGI), MASA syndrome and X-linked
 RT hydrocephalus result from mutations in the L1 gene.";
 RL Nat. Genet. 7:402-407(1994).
 [14]
 RP VARIANTS MASA GLN-210 AND ASN 598.
 RX MEDLINE-95004609; PubMed-7920660;
 RA Vits L., van Camp G., Coucke P., Fransen E., de Boule K.,
 RA Reyniers E., Korn B., Bouska A., Wilson G., Schrander-Stumpel C.,
 RA Winter R.M., Schwartz C., Willems P.J.;
 RT "MASA Syndrome is due to mutations in the neural cell adhesion gene
 RT L1CAM.";
 RL Nat. Genet. 7:408-413(1994).
 [15]
 RP VARIANTS HSAS/MASA S-9; S-121; K-309, F-768; U-94; AND C-1070.
 RX MEDLINE-95282776; PubMed-7762552;
 RA Jouet M., Moncla A., Paterson J., McKeown C., Fryer A., Carpenter N.,
 RA Holmberg E., Wadellius C., Kenwright S.;
 RT "New domains of neural cell adhesion molecule L1 implicated in
 RT X-linked hydrocephalus and MASA syndrome.";
 RL Am. J. Hum. Genet. 56:1304-1314(1995).
 [16]
 RP VARIANTS HSAS/MASA Q-184; Q-210; Y-264; R-452; N-598 AND L-1194.
 RX MEDLINE-96153146; PubMed-8556302;
 RA Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
 RT "CRASH syndrome: clinical spectrum of corpus callosum hypoplasia,
 RT retardation, adducted thumbs, spastic paraparesis and hydrocephalus
 RT due to mutations in one single gene, L1.";
 RL Eur. J. Hum. Genet. 3:273-284(1995).
 [17]
 RP PARALUM.
 RA Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
 RL Eur. J. Hum. Genet. 4:126-126(1996).
 [18]
 RP VARIANTS HSAS/MASA/SPG1 SER-179 AND ARG-370.
 RX MEDLINE-96057511; PubMed-7562969;
 RA Ruiz J.C., Cuppens H., Legius E., Frys J.-P., Glover T., Marynen P.,
 RA Cassiman J.-J.;

"Mutations in L1-CAM in two families with X linked complicated
 RT spastic paraplegia, MASA syndrome, and HSAS";
 RL J. Med. Genet. 32:549-552(1995).
 [19]
 RP VARIANTS HSAS CYS-194 AND LEU-240.
 RX MEDLINE-97083370; PubMed-8929944;
 RA Gu S.-M., Orth U., Veske A., Enders H., Kluender K., Schloesser M.,
 RA Engel W., Schwinger E., Gal A.;
 RT "Five novel mutations in the L1CAM gene in families with X linked
 RT hydrocephalus.";
 RL J. Med. Genet. 33:103-106(1996).
 [20]
 RP VARIANTS HSAS O-184; 439-V-T-443 DEL; C-784 AND 936-G-L-948 DEL.
 RX MEDLINE-97338664; PubMed-9195224;
 RA Macfarlane J.R., Du J.-S., Pepys M.E., Ramsden S., Donnai D.,
 RA Charlton R., Garrett C., Tolmie J., Yates J.R.W., Berry C., Goudie D.,
 RA Monica A., Hunt P., Hodgson S., Jouet M., Kenwright S.;
 RT "Nine novel L1 CAM mutations in families with X-linked
 RT hydrocephalus.";
 RL Hum. Mutat. 9:512-518(1997).
 [21]
 RP VARIANTS HSAS/MASA ASP-691; ARG-698 AND PRO-935.
 RX MEDLINE-98180721; PubMed-9521424;
 RA Du Y.-Z., Srivastava A.K., Schwartz C.E.;
 RT "Multiple exon screening using restriction endonuclease
 RT fingerprinting (REF): detection of six novel mutations in the L1 cell
 RT adhesion molecule (L1CAM) gene.";
 RL Hum. Mutat. 11:232-230(1998).
 [22]
 RP VARIANT CRASH PRO-632.
 RX MEDLINE-98112489; PubMed-9452110;
 RA Vits L., Chitayat D., van Camp G., Holden J.J.A., Fransen E.,
 RA Willems P.J.;
 RT "Evidence for somatic and germline mosaicism in CRASH syndrome.";
 RL Hum. Mutat. Suppl. 1:S284-S287(1998).
 [23]
 RP VARIANTS HSAS/MASA THR-219; ARG-335; CYS-386; CYS-473 AND LEU-1224.
 RX MEDLINE-98415726; PubMed-9744477;
 RA Saugier-Verber P., Martin C., le Meur N., Lyonnet S., Munnich A.,
 RA David A., Henocq A., Heron D., Jonveaux P., Odent S., Manouvrier S.,
 RA Monica A., Moirichon N., Philip N., Saïde D., Tosi M., Frebourg T.;
 RT "Identification of novel L1CAM mutations using fluorescence-assisted
 RT mismatch analysis.";
 RL Hum. Mutat. 13:259-366(1998).
 CC -!- FUNCTION: CELL ADHESION MOLECULE WITH AN IMPORTANT ROLE IN THE
 CC DEVELOPMENT OF THE NERVOUS SYSTEM. INVOLVED IN NEURON-NEURON
 CC ADHESION. NEURITE FASCICULATION, OUTGROWTH OF NEURITES, ETC. BINDS
 CC TO AXONIN ON NEURONS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -!- DISEASE: DEFECTS IN L1CAM ARE THE CAUSE OF THREE X-LINKED
 CC SYNDROMES 1: HYDROCEPHALUS OWING TO STENOSIS OF THE AQUEDUCT OF
 CC SYLVIVUS (HSAS) CHARACTERIZED BY MENTAL RETARDATION AND ENLARGED

Query Match

Best Local Similarity 16.2%; Score 96.5; DB 1; Length 1257;
 : Matches 35; Conservative 12; Mismatches 40; Indels 21; Gaps 4;
 QY 1 AONTARIGEPLVLK--KGAPKPPQRLKLTGRTAEKVKISPOGGPMDSVARVLP 58
 DB 249 SSHLVAGGOLPLVLECAEGFP-----TPTIKWLRLSPGMPADRVYQNH 293
 QY 59 NGSFLPAVGITODEGIFRCQAMNNGKETSRYRVYQIP---GRPE 103
 DB 294 NKTLLQLKVGEDDGEYRCIAENSLG-SARHAYTVVEAAPYWLHKPQ 340

RESULT 11

AXOL_RAT
 ID AXOL_RAT
 AC P22063;
 DT 01-AUG-1991 (rel. 19, Created)

pfam: PF00052; laminin_B; 2.
 DR Pfam: PF00053; laminin_EGF; 5.
 DR Pfam: PF00057; ldl_receptor; 3.
 DR ProDom: PD003031; laminin_B; 2.
 DR SMART: SM00180; EGF_Lam; 5.
 DR SMART: SM00410; IG_Like; 5.
 DR SMART: SM00408; IG2; 11.
 DR SMART: SM00192; IG2; 3.
 DR SMART: SM00281; Lamb; 2.
 DR SMART: SM00018; P; PARTIAL.
 DR PROSITE: PS00022; EGF_1; 4.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01209; LDL_RA_1; 3.
 DR PROSITE: PS00068; LDL_RA_2; 3.
 DR PROSITE: PS01248; LAMININ_TYE_EGF; 7.
 DP proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;
 KW Basemem matrix; Extracellular matrix; Alternative splicing;
 KW laminin EGF-like domain.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 2481
 FT DOMAIN 19 148 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 149 188 LDL-RECEPTOR CLASS A 1.
 FT DOMAIN 190 229 LDL-RECEPTOR CLASS A 2.
 FT DOMAIN 233 272 LDL-RECEPTOR CLASS A 3.
 FT DOMAIN 273 363 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 384 431 LAMININ EGF-LIKE 1 (INCOMPLETE).
 FT DOMAIN 432 441 LAMININ EGF-LIKE 2 (N-TERMINAL).
 FT DOMAIN 442 632 LAMININ DOMAIN IV 1.
 FT DOMAIN 633 665 LAMININ EGF-LIKE 2 (C-TERMINAL).
 FT DOMAIN 673 719 LAMININ EGF-LIKE 3 (INCOMPLETE).
 FT DOMAIN 720 729 LAMININ EGF-LIKE 4 (N-TERMINAL).
 FT DOMAIN 730 920 LAMININ DOMAIN IV 2.
 FT DOMAIN 921 953 LAMININ EGF-LIKE 4 (C-TERMINAL).
 FT DOMAIN 954 1003 LAMININ EGF-LIKE 5.
 FT DOMAIN 1010 1059 LAMININ EGF-LIKE 6.
 FT DOMAIN 1060 1110 LAMININ EGF-LIKE 7.
 FT DOMAIN 1133 1216 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 1227 1311 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 1318 1403 IG-LIKE C2-TYPE DOMAIN 5.
 FT DOMAIN 1415 1499 IG-LIKE C2-TYPE DOMAIN 6.
 FT DOMAIN 1507 1592 IG-LIKE C2-TYPE DOMAIN 7.
 FT DOMAIN 1598 1682 IG-LIKE C2-TYPE DOMAIN 8.
 FT DOMAIN 1699 1785 IG-LIKE C2-TYPE DOMAIN 9.
 FT DOMAIN 1794 1878 IG-LIKE C2-TYPE DOMAIN 10.
 FT DOMAIN 1887 1976 IG-LIKE C2-TYPE DOMAIN 11.
 FT DOMAIN 1977 2070 IG-LIKE C2-TYPE DOMAIN 12.
 FT DOMAIN 2078 2162 IG-LIKE C2-TYPE DOMAIN 13.
 FT DOMAIN 2174 2260 IG-LIKE C2-TYPE DOMAIN 14.
 FT DOMAIN 2267 2347 IG-LIKE C2-TYPE DOMAIN 15.
 FT DOMAIN 2354 2436 IG-LIKE C2-TYPE DOMAIN 16.
 FT DISULFID 66 114 BY SIMILARITY.
 FT DISULFID 149 161 BY SIMILARITY.
 FT DISULFID 156 174 BY SIMILARITY.
 FT DISULFID 168 184 BY SIMILARITY.
 FT DISULFID 190 202 BY SIMILARITY.
 FT DISULFID 197 215 BY SIMILARITY.
 FT DISULFID 209 224 BY SIMILARITY.
 FT DISULFID 233 246 BY SIMILARITY.
 FT DISULFID 240 259 BY SIMILARITY.
 FT DISULFID 253 268 BY SIMILARITY.
 FT DISULFID 954 963 BY SIMILARITY.
 FT DISULFID 956 970 BY SIMILARITY.
 FT DISULFID 973 982 BY SIMILARITY.
 FT DISULFID 985 1001 BY SIMILARITY.
 FT DISULFID 1010 1020 BY SIMILARITY.
 FT DISULFID 1012 1026 BY SIMILARITY.
 FT DISULFID 1029 1038 BY SIMILARITY.
 FT DISULFID 1041 1057 BY SIMILARITY.
 FT DISULFID 1060 1068 BY SIMILARITY.
 FT DISULFID 1062 1078 BY SIMILARITY.
 FT DISULFID 1081 1090 BY SIMILARITY.
 FT DISULFID 1093 1108 BY SIMILARITY.
 FT DISULFID 1151 1199 BY SIMILARITY.
 Query Match 16.1%; Score 95.5; DB 1; Length 2481;
 Best local similarity 25.4%; Pred. No. 0.092;
 Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;
 QY 8 IGEPVLVKCK--CAPKKPPORLEWKLNTGRTAEKVKLSPOGG-PWDSVARVLPNGSLFL 64
 DL 1989 VGEPLQVKCEAFGAPGDPEVEWLDHPG-----PERGDLDPDFKPTISEQFIRH 2039
 QY 65 PAVCIODEGIFRCQAMNRNGKTKSNY-----RVRVYQIPG-----KPEIVDS 107
 DL 2040 PNVGLGNAGVYCKGSSAHATATKNIYIEVPEPSRIATVSLGSSQWFQDQGEGLICT 2099
 QY 108 AS 109
 DL 2100 AT 2101
 RESULT 13
 GAC1_HUMAN STANDARD; PRT: 713 AA.
 ID GAC1_HUMAN
 AC 075325; 2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glioma amplified on chromosome 1 protein precursor.
 GN GAC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC TISSUE-Glial tumor;
 RX MEDLINE-98324709; PubMed-9662332;
 RA Malfroy B., Alnoide A., Zhu X.X., Vogt N., Tyagi R., Muleris M.,
 RA Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;
 RT "GAC1, a new member of the leucine-rich repeat superfamily on
 RT chromosome band 1q32.1, is amplified and overexpressed in malignant
 RT gliomas.";
 KL Oncogene 16:2947-3002(1998).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential)
 CC -!- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: AF030435; AAC39792.1;
 DR MIM: 605492;
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_c2.

Search completed: May 30, 2004, 15:55:37
Job time : 8.55856 secs

Search completed: May 30, 2004, 15:55:37
Job time : 8.55856 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Computer Aid, Inc.

OM protein - protein search, using sw model

```
Run on: May 30, 2003, 15:47:46 ; Search time: 25.982 seconds
      (without alignments)
      888.203 Million cell updates/sec
```

Title: US-09-872-185R-1

Perfect score:

Sequence: I AQNITARIIEPIVKKKQAP VEVYQIP;FPELVISASELT 112

Scoring table: BLOSUM62

Gapop 10 0 : Gapext 0 5

Searched: 671580 sigs. 206047115 residues

Category	Number of hits satisfying chosen parameters
Total	677580

Minimum DR seed length: 0

Minimum	DB seq	length:	0
Maximum	DB seq	length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0*
Maximum Match 100%Max Inquiry Mail 1008
Listing first 45 summaries

Database :

SPTREMBL 21:★

```

DEFINITION_21:
1:  sp archea: *

```

1: sp_at:clad:*

2: sp_bacteria:*

3: sp_lung1: *

4: sp_human: *human*
5: sp_invertebrate: *

6: sp_mammal.*

6: sp_mdm1:*

8: sp. organellae

8: sp_Organelle:
9: sp_Organelle.*

```
9: sp_pragc: *
10: sp_plant: *
```

10: sp_plant: *

11: sp_rodent: *

11: sp_rotent:★
12: sp_virus:★12: sp_verts: *
13: sp_verts: *

```
13: sp_verlebrate:*
14: sp_unclassified:*
```

```
14: sp_unclassified
15: sp_unknown
```

15: sp_rv415:

```
16: sp_base = 0
17: sp_base = 0
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No	Score	%		Length	PR	ID	Description
		Query Match					
1	516.5	86.8		402	11	Q35444	Q35444 mus musculus
2	118	19.8		32	6	Q4TR01	Q4TR01 bos taurus
3	114	19.2		1445	11	Q63155	Q63155 ratius norv
4	101	17.0		1212	5	Q95T60	Q95T60 drosophila
5	101	17.0		2016	5	Q9NBA1	Q9NBA1 drosophila
6	101	17.0		2016	5	Q9V4J9	Q9V4J9 drosophila
7	99	16.6		628	4	Q4K180	Q4K180 homo sapien
8	96.5	16.2		1248	6	Q9XT41	Q9XT41 cercopithe
9	96	16.1		1018	6	Q28106	Q28106 bos taurus
10	96	15.1		1344	11	Q9ZT14	Q9ZT14 mus musculus
11	95.5	16.1		2295	5	Q9XT62	Q9XT62 caenorhabdi
12	95.5	16.1		2482	5	Q38253	Q38253 caenorhabdi
13	95.5	16.1		3375	5	Q9XT15	Q9XT15 caenorhabdi
14	95	16.0		182	4	Q15402	Q15402 homo sapien
15	95	16.0		1034	4	Q96MS0	Q96MS0 homo sapien
16	95	16.0		1427	13	Q91562	Q91562 xenopus lae

DI 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-JUN-2002 (TrEMBLrel. 15, Last sequence update)
 DE Dscam precursor
 GN DSCAM OR CG17800.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 II 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20348742; PubMed=10872653;
 RA Schumacker D., Clemens J.C., Shu H., Worby C.A., Xiao J., Muda M.,
 RA Dixon J.E., Zipursky S.J.;
 RT "Drosophila Dscam is an Axon Guidance Receptor Exhibiting
 RT Extraordinary Molecular Diversity";
 RT Cell 101:671-684(2000).
 DR FBL: AF260530; AAF71926.1; -;
 DR HSPF: P40189; IBDQ.
 DR FlyBase; FBgn0033159; Dscam.
 DR InterPro: IPR000267; Asp/Glutamase.
 DR InterPro: IPR003961; FN.III.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; Ig; 9.
 DR SMART: SM00060; FN3; 6.
 DR SMART: SM00408; IgC2; 7.
 DR SMART: SM00410; Ig_Like; 3.
 DR PROSITE: PS00144; ASN_GLN_ASE_1; UNKNOWN_1.
 KW Immunoglobulin domain; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 SQ SEQUENCE 2016 AA; 222124 MW; 95CF95488FZAD36C CRC64;

 Query Match 17.0%; Score 101; DB 5; Length 2016;
 Best Local Similarity 29.2%; Pred. No. 0.06;
 Matches 33; Conservative 18; Mismatches 48; Indels 14; Gaps 5;

 QY 2 ONITARIGEPVLVKCGAPKPPQPLEWKLNTGRTEAWKVLSPGGGPPWSVARVLPNG- 60
 DB 819 RNOTARGEPVAVLQCEAKGKPGIGL-WNNMNM-----LDPKNDNRYYTIREEILSTGV 871

 QY 61 -SLPLPAVGIGDEGIFRCQAMNNGKTKSNRYRVYQIPGKP---EIVDSA 108
 DB 872 MSSLSIKRTSRSDSALFTCVATNAFGSDAS-INMIVQEVPEMPYALKVLKDS 923

 RESULT 6
 ID Q9V4J9 PRELIMINARY; PRT; 2016 AA.
 AC Q9V4J9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE CG17800 protein.
 GN DSCAM OR CG17800.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 II 111
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceinlier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon P.C., Rogers Y.-H., Blazei R.G., Champe M., Pfeiffer B.B.,
 RA Wan K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew P.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 RA Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.R., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evans J.C., Ferraz C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Finkler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez H., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kocira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liao X., Mattei P., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Palazzolo M., Pittman G.S., Pan S., Poulard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidon-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Stadling A.C., Stapleton M., Strong P., Sun E.,
 RA Svirkas P., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Williams S.M., Woodage T., Worley K.C., Wu L., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang C., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 PL Science 287:2185-2195(2000).
 DR FBL: AE003841; AAF59271.1; -;
 DR HSPF: P40189; IBDQ.
 DR FlyBase; FBgn0033159; Dscam.
 DR InterPro: IPR000267; Asp/Glutamase.
 DR InterPro: IPR003961; FN.III.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00041; fn3; 6.
 DR Pfam: PF00047; Ig; 9.
 DR SMART: SM00060; FN3; 6.
 DR SMART: SM00408; IgC2; 7.
 DR SMART: SM00410; Ig_Like; 2.
 DR PROSITE: PS00144; ASN_GLN_ASE_1; UNKNOWN_1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 2016 AA; 222109 MW; 64A8DE3BB7ED0AB0 CRC64;

 Query Match 17.0%; Score 101; DB 5; Length 2016;
 Best Local Similarity 29.2%; Pred. No. 0.06;
 Matches 33; Conservative 18; Mismatches 48; Indels 14; Gaps 5;

 QY 2 ONITARIGEPVLVKCGAPKPPQPLEWKLNTGRTEAWKVLSPGGGPPWSVARVLPNG- 60
 DB 819 RNOTARGEPVAVLQCEAKGKPGIGL-WNNMNM-----LDPKNDNRYYTIREEILSTGV 871

 QY 61 -SLPLPAVGIGDEGIFRCQAMNNGKTKSNRYRVYQIPGKP---EIVDSA 108
 DB 872 MSSLSIKRTSRSDSALFTCVATNAFGSDAS-INMIVQEVPEMPYALKVLKDS 923

 RESULT 7
 ID Q9BTNO PRELIMINARY; PRT; 628 AA.
 AC Q9BTNO;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 66.3 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg K.;
RL Submitted (FEB-2001) to the EMBL/GenBank/Trna databases
DR EMBL: BC003578; AA003578.1;
DR InterPro: IPR003961; FN 111;
DR InterPro: IPR003598; Iq;
DR InterPro: IPR003598; Iq; C2;
DR InterPro: IPR003006; Iq; MHC;
DR InterPro: IPR001611; LRR;
DR InterPro: IPR000483; LRR; C-term;
DR InterPro: IPR000372; LRR; N-term;
DR InterPro: IPR003591; LRR; TYP;
DR Pfam: PF00041; fn3; 1;
DR Pfam: PF00047; Iq; 1;
DR Pfam: PF00560; LRR; 6;
DR Pfam: PF01463; LRRCT; 1;
DR PRINTS: PR00019; LEURICIRPT;
DR SMART: SM00060; FN3; 1;
DR SMART: SM00409; Iq; 1;
DR SMART: SM00408; Iq; C2; 1;
DR SMART: SM00082; LRRCT; 1;
DR SMART: SM00013; LRRNT; 1;
DR SMART: SM00369; LRR; TYP; 6;
KW Hypothetical protein; Immunoglobulin domain;
SQ SEQUENCE 628 AA: 66259 MW: 6835A5056AHC5FC GRC64;

Query Match 16.6%; Score 99; DB 4; Length 628;
Best Local Similarity 31.5%; Pred. No. 0.024;
Matches 34; Conservative 12; Mismatches 44; Indels 28; Gaps 6;

QY 9 GEPVLVLCGKAPKKPPORLEWKLNTGTEAWKVLSPQ--GGIWDSSVARVLNGSLFLP 65
DB 310 GRPAALRCR-AGVDFEVRVW ----- VSPQGRLLGN -SSRAKAFNGTLELL 354
QY 66 AVGLQDEGIFRCQAMNENKETSNNYRVVYQI--- PKPEIVDSAS 109
DB 355 VIEGIDGGIFPTCAANAAGEATAA -VELIVGPPPPQIANSIS 396

RESULT 8
QYX141
ID Q9X141 PRELIMINARY; PRI: 1248 AA;
AC Q9X141;
DI 01-NOV-1999 (TREMBLrel. 12; Created)
DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19; Last annotation update)
DE Neural cell adhesion molecule L1.
GN L1CAM.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea.
OX NCBI_TaxID=9514;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Fackell U., Schroder J., Resstet B., Veske A., Gal A.;
RL "Spectrum and detection rate of L1CAM mutations in isolated and familial cases with clinically suspected L1-disease.";
RL Am. J. Med. Genet. 92:40-46(2000).
DR EMBL: AF129167; AA028610.1;
DR HSP: F20241; 1CFB;
DR InterPro: IPR004961; FN 111;
DR InterPro: IPR003598; Iq; C2;
DR InterPro: IPR003600; Iq; like;
DR InterPro: IPR004006; Iq; MHC;
DR Pfam: PF00041; fn3; 5;
DR Pfam: PF00047; Iq; 6;
DR SMART: SM00060; FN3; 4;

Query Match 16.1%; Score 96; DB 6; Length 1018;
Best Local Similarity 26.4%; Pred. No. 0.091;
Matches 29; Conservative 26; Mismatches 35; Indels 20; Gaps 6;

QY 2 GNITAPIGEPHVLKCKGAKHPFPPELEWKLNTGTEAWKVLSPQGGIWDSSVARVLNPS 61
DB 249 KIVYALMGNVITPEFALGNIVPPIHWP ----- KVIQPM PSTALEISCA 294
QY 62 LEFLAVGLQDEGIFRCQAMNENKETSNNYRVVYQIWKPEIVDSASE 110
DB 244 VIKTNIQLEFGVYFCEAENNPEN-- KIGARIY VQALPEWVEHND 339

RESULT 10
QYX214
ID QYX214 PRELIMINARY; PRI: 1344 AA;
AC QYX214;

```


Fri May 30 17:16:34 2003

```

DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000034; Laminin_B.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR000519; P_trefoil.
DR Pfam: PF00047; Ig_16.
DR Pfam: PF00052; laminin_B; 2.
DR Pfam: PF00053; laminin_EGF; 5.
DR Pfam: PF00054; laminin_G; 2.
DR Pfam: PF00057; ldl_recept_a; 3.
DR SMART: SM00410; IG_Like; 5.
DR SMART: SM00281; Lamb; 2.
DR SMART: SM00192; LDLA; 3.
DR SMART: SM00018; PD; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_4.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 7.
DR PROSITE: PS01209; LDLA_1; 3.
DR PROSITE: PS00068; LDLA_2; 4.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Alternative splicing; EGF-like domain; Glycoprotein;
KW Hypothetical protein; immunoglobulin domain; Laminin EGF like domain;
KW Repeat.
SQ SEQUENCE 2482 AA: 270944 MW: 824006 FDR64070 CRC64;
Query Match 16.1%; Score 95.5; DH 5; Length 2482;
Best local Similarity 25.4%; Pred. No. 0.31;
Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;
QY 8 IGEPVLKCK-GAPKKPPQRLKLTGRTENMKVLSPOGSG-PWDSVAVPLPNSGLFL 64
DB 1990 VGEPLQVQKCAFCAQDPPEVFWLHDPG-----PERGDLDDPKPVTISEQPIRH 2040
QY 65 PAVGIQDEGIFRCQAMNNGKTKSNY-----KRVVYQIPG-----KPEIVDS 107
DB 2041 PNVLGNAGVYTCCKSSAHATATKNLYTEVVEPSIATVSLGGSSQWFDGEGKGLICT 2100
QY 108 AS 109
DB 2101 AT 2102
RESULT 13
Q9XT15 PRELIMINARY; PRT; 3375 AA.
AC Q9XT15;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE UNC-52 protein.
CN UNC-52 OR ZC101.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA Percy C., Barnes C.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS: FOUR FORMS: ISOFORM ZC101.2A (Q9XT15),
CC ISOFORM ZC101.2B (Q1B2G1), ISOFORM ZC101.2C (Q9XT02) AND ISOFORM
CC ZC101.2E (SHOWN HERE); MAY BE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL: Z93395; CAB07708.1;
DR EMBL: Z93375; CAB07708.1; JOINED.
DR EMBL: Z93375; CAB07569.1;
DR EMBL: Z93395; CAB07569.1; JOINED.
DR HSP: P01130; I1DR.
DR wormPep, ZC101.2E; GE18424.

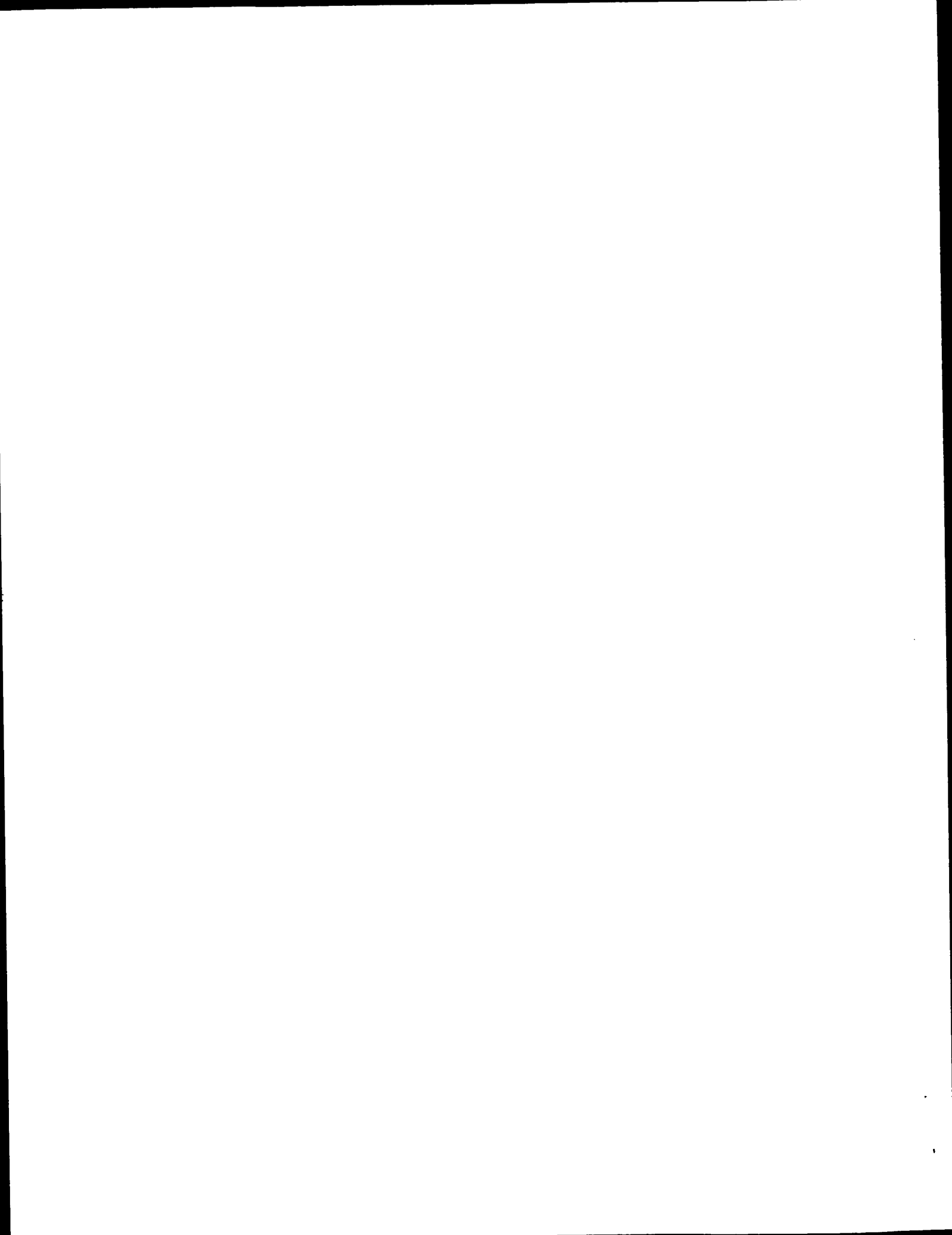
```

```

DR InterPro: IPR000561; EGF_Like.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000034; Laminin_B.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR000519; P_trefoil.
DR Pfam: PF00047; Ig_17.
DR Pfam: PF00052; laminin_B; 2.
DR Pfam: PF00053; laminin_EGF; 5.
DR Pfam: PF00054; laminin_G; 2.
DR Pfam: PF00057; ldl_recept_a; 3.
DR PRINTS: PR00261; LDLRECEPTOR.
DR ProDom: PD003031; Laminin_B; 2.
DR SMART: SM00180; EGF_Lam; 5.
DR SMART: SM00001; EGF_Like; 3.
DR SMART: SM00408; IGC2; 12.
DR SMART: SM00281; IG_Like; 5.
DR SMART: SM00281; Lamb; 2.
DR SMART: SM00282; Lamb; 3.
DR SMART: SM00192; LDLA; 3.
DR SMART: SM00018; PD; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 7.
DR PROSITE: PS01209; LDLA_1; 3.
DR PROSITE: PS00068; LDLA_2; 4.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Alternative splicing; EGF-like domain; Glycoprotein;
KW Hypothetical protein; immunoglobulin domain; Laminin EGF like domain;
KW Repeat.
SQ SEQUENCE 3375 AA: 369049 MW: 1AA418B34E5D67AA CRC64;
Query Match 16.1%; Score 95.5; DH 5; Length 3375;
Best local Similarity 25.4%; Pred. No. 0.45;
Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;
QY 8 IGEPVLKCK-GAPKKPPQRLKLTGRTENMKVLSPOGSG-PWDSVAVPLPNSGLFL 64
DB 1990 VGEPLQVQKCAFCAQDPPEVFWLHDPG-----PERGDLDDPKPVTISEQPIRH 2040
QY 65 PAVGIQDEGIFRCQAMNNGKTKSNY-----KRVVYQIPG-----KPEIVDS 107
DB 2041 PNVLGNAGVYTCCKSSAHATATKNLYTEVVEPSIATVSLGGSSQWFDGEGKGLICT 2100
QY 108 AS 109
DB 2101 AT 2102
RESULT 14
Q15402 PRELIMINARY; PRT; 182 AA.
AC Q15402;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE pregnancy-specific protein (fragment).
CN PSC1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS.
RX MEDLINE-96147764; PubMed-2392228;
RX Rorjidan J., Tease L.A., Barnes W., Chan W.Y.;
RX "Expression of the pregnancy-specific beta 1-glycoprotein genes in
RX human testis.";

```

Search completed: May 30, 2003, 15:57:28
Job time : 28.982 secs



OM protein - protein search, using sw model

Run on: May 30, 2003, 14:53:06 : Search time 75.5225 seconds
(without alignments)
585,775 Million cell updates/sec

Title: US-09-872-185b-2

Perfect score: 1766

Sequence: 1 AUNTAPR:EFIVEK:KAP

Scoring table: BIOSUM62

Gapop 10 0 0 Gapovt 0 5

Searched: 508470 seqs, 14545620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 40 summaries

Database: A_Geneseq_101002.*

1	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1980.DAT.*
2	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1981.DAT.*
3	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1982.DAT.*
4	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1983.DAT.*
5	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1984.DAT.*
6	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1985.DAT.*
7	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1986.DAT.*
8	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1987.DAT.*
9	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1988.DAT.*
10	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1989.DAT.*
11	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1990.DAT.*
12	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1991.DAT.*
13	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1992.DAT.*
14	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1993.DAT.*
15	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1994.DAT.*
16	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1995.DAT.*
17	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1996.DAT.*
18	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1997.DAT.*
19	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1998.DAT.*
20	/SIDS2/gqdata/geneseq/geneseq-emb1/AA1999.DAT.*
21	/SIDS2/gqdata/geneseq/geneseq-emb1/AA2000.DAT.*
22	/SIDS2/gqdata/geneseq/geneseq-emb1/AA2001.DAT.*
23	/SIDS2/gqdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DR	ID	Description
1	1766	100.0	404	23	AAH81925	Extracorporeal circulation material receptor protein.
2	1766	100.0	404	23	AAH81925	Human RAGE protein
3	1762	99.8	404	23	AAH81925	Human RAGE protein
4	1762	99.8	404	23	AAH81925	Human receptor for
5	1762	99.8	404	23	AAH81925	Human receptor for
6	1701	96.3	318	18	AAW4200	Human mature recep
7	1701	96.3	318	18	AAW4200	Human RAGE polypep
8	1701	96.3	318	18	AAW4200	Human soluble rece
9	1701	96.3	318	18	AAW4200	Human RAGE polypep
10	1696	96.0	339	23	AAH81925	Human RAGE protein

11	1463	82.8	416	23	AAH81925	Cow receptor for a
12	1463	82.8	416	23	AAH81925	Bovine receptor fo
13	1383.5	78.3	403	23	AAH81925	Mouse receptor for
14	1383.5	78.3	403	23	AAH81925	Mouse receptor fo
15	595	33.7	112	23	AAH81925	Human RAGE protein
16	220.5	12.5	582	23	AAH81925	Human lung tumour
17	220.5	12.5	582	23	AAH81925	Clone #18964 of lu
18	218.5	12.4	583	19	AAW4200	Human activated le
19	218.5	12.4	583	19	AAW4200	Activated leukocyt
20	217.5	12.3	646	20	AAH81925	Human MUC18 amino
21	217.5	12.3	646	20	AAH81925	Human MUC18 amino
22	216.5	12.3	646	20	AAH81925	Human MUC18 amino
23	214.5	12.1	570	17	AAH81925	Stem cell marker H
24	213.5	12.1	574	21	AAH81925	Amino acid sequenc
25	212.5	12.0	753	20	AAH81925	Human T85 protein.
26	210	11.9	1496	20	AAH81925	Melanoma associate
27	210	11.9	1496	20	AAH81925	Human p53 target m
28	210	11.9	1496	20	AAH81925	Human peroxidasin
29	209	11.8	1449	20	AAH81925	Human p90gl protein
30	209	11.8	1449	20	AAH81925	Human p90gl protein
31	208.5	11.8	583	17	AAH81925	Stem cell marker H
32	204.5	11.6	511	16	AAH81925	BGP (1-314)/CEA (4
33	204	11.6	5635	23	AAH81925	Novel human protei
34	203	11.5	1033	22	AAH81925	Proscophila melano
35	200.5	11.4	509	16	AAH81925	BGP (1-314)/CEA (4
36	200	11.3	493	16	AAH81925	BGP (1-314)/CEA (4
37	197.5	11.2	404	22	AAH81925	Human NOV12 protei
38	196.5	11.1	206	23	AAH81925	Human ovarian anti
39	196.5	11.1	782	23	AAH81925	Human polypeptide
40	196.5	11.1	844	23	AAH81925	Human polypeptide
41	194.5	11.0	898	22	AAH81925	Novel human diagno
42	194	10.9	1483	22	AAH81925	Novel human diagno
43	191.5	10.8	628	9	AAH81925	Carcinoblastic a
44	190.5	10.8	679	23	AAH81925	Human carcino-emb
45	190.5	10.8	698	16	AAH81925	Carcinoembryonic a

ALIGNMENTS

RESULT 1

AAH81925

ID: AAH81925 standard; protein; 404 AA.

AC: AAH81925;

XX: 15 JUN 2001 (first entry)

DE: Extracorporeal circulation material receptor protein.

XX: Extracorporeal circulation; carbonyl stress product; receptor;

KW: diabetes; vascular lesion; excretory dysfunction.

XX: Unidentified.

OS: Unidentified.

PN: W0200118060-AA.

XX: 15-MAR-2001.

PD: 08-SEP-2000; 2000CW-3P06172.

PF: 08-SEP-1999; 99JP-0254463.

PP: (COFA) TORAY INC.

XX: Shimizu S, Kikuta M, Akiyama H, Usui M.

XX: WPI: 2001-290314/30.

XX: Material for extracorporeal circulation, applicable in selective

PI: elimination of diabetic complication factors such as carbonyl stress

PT: products caused by abnormally promoted carbonyl stress from excretory

PT: dysfunction in vascular lesions.

XX PS Claim 1: Page 31 32: 36pp; Japanese.

XX CC the present invention describes a material for extracorporeal circulation

XX CC which is made from a water insoluble carrier immobilized with a protein

XX CC having the sequence shown here. The materials of the invention, including

XX CC adsorbents, are for extracorporeal circulation, which are applicable in

XX CC the selective elimination of diabetic complication factors from a body

XX CC fluid, and are therefore useful in treating vascular lesions like

XX CC arteriosclerosis due to carbonyl stress products caused by abnormally

XX CC promoted carbonyl stress from excretory dysfunction.

XX SC Sequence 404 AA:

Query Match 100.0%; Score 1766; DB 22; Length 404;

Best Local Similarity 100.0%; Pred. No. 1.2e-124;

Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGNITARIGELVVKCKGAPKPPQPLFWKLNIGRLKAWKVSQGGGWDVSARVLPNG 60

DB 23 AGNITARIGELVVKCKGAPKPPQPLFWKLNIGRLKAWKVSQGGGWDVSARVLPNG 82

QY 61 SLFLPVGILQDEGIFRCQAMNNGKETKSNRYRVVYQIPCKPELVDSASELTACVKNKG 120

DB 83 SLFLPVGILQDEGIFRCQAMNNGKETKSNRYRVVYQIPCKPELVDSASELTACVKNKG 142

QY 121 TCVSFCSPGACILSWHLGCKPLVINEKGVSVKQETRRHPETGLFTQSELMTVPARGGDP 180

DB 143 TCVSFCSPGACILSWHLGCKPLVINEKGVSVKQETRRHPETGLFTQSELMTVPARGGDP 202

QY 181 RPTFSCSPGLPRHRLPTAPIQPRVWPEVPLEEVOLVVEPEGAVAPGTVTLTCEVP 240

DB 203 RPTFSCSPGLPRHRLPTAPIQPRVWPEVPLEEVOLVVEPEGAVAPGTVTLTCEVP 262

QY 241 AOPSQIHWKKGVPPLPSPVLILPEIGPDQCTYSCVATSSHGQPSRAVSISIE 300

DB 263 AOPSQIHWKKGVPPLPSPVLILPEIGPDQCTYSCVATSSHGQPSRAVSISIE 322

QY 301 PGEEGPTAGSVGGSGLTALALGILGGLTA 332

DB 323 PGEEGPTAGSVGGSGLTALALGILGGLTA 354

RESULT 2

AAM48745

ID AAM48745 standard; protein; 404 AA.

XX AC AAM48745;

XX DT 02-APR-2002 (first entry)

XX DE Human RAGE protein SEQ ID NO 1.

XX KW Human: RAGE; receptor for advanced glycosylated endproduct; receptor;

XX KW antidiabetic; neuroprotective; cytostatic; antiinflammatory; vasoregic;

XX KW nephrotropic; dermatological; antarteriosclerotic; nootropic; diabetes;

XX KW Alzheimer's disease; cancer; inflammation; kidney failure;

XX KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.

XX OS Homo sapiens.

XX FN W0200192892-A2.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-0517447.

XX PR 30-MAY-2000; 2000US-207442P.

XX PR 05-MAR-2001; 2001US-0799152.

XX PA (TRAN-) TRANS TECH PHARMA.

XX PR Shabbaz M.

XX WP: 2002 114372/15.

XX CC detecting a receptor for advanced glycosylated endproducts (RAGE)

XX CC modulators, for treating e.g., cancer, diabetes or inflammation,

XX CC comprises measuring the amount of bound anti RAGE antibody

XX CC Claim 1: Fig 2; 49pp; English.

XX CC The invention relates to detecting receptor for advanced glycosylated

XX CC endproducts (RAGE) modulators comprises determining the amount of RAGE

XX CC protein or its fragment bound to the pre-adsorbed bound by measuring the

XX CC amount of anti-RAGE antibody bound to the solid surface. The method is

XX CC useful for rapid, high-throughput identification of compounds that

XX CC modulate RAGE. The compounds are useful for treating symptoms of diabetes

XX CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's

XX CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis

XX CC or inflammatory lupus nephritis, erectile dysfunction and

XX CC atherosclerosis.

XX SC Sequence 404 AA:

Query Match 100.0%; Score 1766; DB 24; Length 404;

Best Local Similarity 100.0%; Pred. No. 1.2e-124;

Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGNITARIGELVVKCKGAPKPPQPLFWKLNIGRLKAWKVSQGGGWDVSARVLPNG 60

DB 23 AGNITARIGELVVKCKGAPKPPQPLFWKLNIGRLKAWKVSQGGGWDVSARVLPNG 82

QY 61 SLFLPVGILQDEGIFRCQAMNNGKETKSNRYRVVYQIPCKPELVDSASELTACVKNKG 120

DB 83 SLFLPVGILQDEGIFRCQAMNNGKETKSNRYRVVYQIPCKPELVDSASELTACVKNKG 142

QY 121 TCVSFCSPGACILSWHLGCKPLVINEKGVSVKQETRRHPETGLFTQSELMTVPARGGDP 180

DB 143 TCVSFCSPGACILSWHLGCKPLVINEKGVSVKQETRRHPETGLFTQSELMTVPARGGDP 202

QY 181 RPTFSCSPGLPRHRLPTAPIQPRVWPEVPLEEVOLVVEPEGAVAPGTVTLTCEVP 240

DB 203 RPTFSCSPGLPRHRLPTAPIQPRVWPEVPLEEVOLVVEPEGAVAPGTVTLTCEVP 262

QY 241 AOPSQIHWKKGVPPLPSPVLILPEIGPDQCTYSCVATSSHGQPSRAVSISIE 300

DB 263 AOPSQIHWKKGVPPLPSPVLILPEIGPDQCTYSCVATSSHGQPSRAVSISIE 322

QY 301 PGEEGPTAGSVGGSGLTALALGILGGLTA 332

DB 323 PGEEGPTAGSVGGSGLTALALGILGGLTA 354

RESULT 4

AAY52140

ID AAY52140 standard; protein; 342 AA.

XX AC AAY52140;

XX DT 28-JAN-2000 (first entry)

XX DE Human Receptor to AGE (RAGE); amino acid sequence.

XX KW Soluble receptor for advanced glycation endproducts; RAGE; tumour;

XX KW invasion; metastasis; amphotericin; neuron; inhibit; therapy.

XX OS Homo sapiens.

XX FN W09954485-A1.

XX PD 28-OCT-1999.

XX PF 16-APR-1999; 99WO-US08427.

XX PR 17-APR-1998; 98US-0062465.

XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX PA Schmidt AM, Stern D;
 XX PI WPI: 2000 01260/01
 XX DR Inhibiting tumour invasion or spreading by administration of soluble
 PT receptor for advanced glycation endproducts
 XX PI
 XX PS Disclosure: Page 10-11; 88pp; English.
 XX
 CC This is the amino acid sequence of the human soluble Receptor for
 CC Advanced Glycation Endproducts (RAGE). RAGE interacts with a range of
 CC physiologically and pathophysiologically relevant ligands when
 CC considering tumour invasion in normal developing neurons. RAGE
 CC colocalizes with amyloid in which is a matrix associated polypeptide.
 CC The expression of both RAGE and amyloid increases after birth, but
 CC both have increased expression in tumours. RAGE polypeptides
 CC AN52132-Y52135 are used in the invention in a method for inhibiting
 CC tumour invasion and metastasis. The method involves inhibiting tumour
 CC invasion and metastasis via administration of a therapeutically effective
 CC amount of the pharmaceutical composition containing a RAGE polypeptide.
 CC The invention also relates to a method for evaluating the ability of an
 CC agent to inhibit tumour invasion in a local cellular environment. RAGE
 CC can be administered to a patient in a pharmaceutically acceptable
 CC carrier.
 XX
 XX SQ Sequence 332 AA;
 Query Match: 99.8%, Score 1763, DB 21, Length 332,
 Best Local Similarity 99.7%, Pred. No. 1.5e-124;
 Matches 331; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQNITARIGEPVLVKCGAPKKPTQPIENKLTGRTTEANKVLSPOGGGPDVSARVLPNG 60
 DB 1 AQNITARIGEPVLVKCGAPKKPTQPIENKLTGRTTEANKVLSPOGGGPDVSARVLPNG 60
 QY 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVQIPCKPIVDASSELTAGVPNKVG 120
 DB 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVQIPCKPIVDASSELTAGVPNKVG 120
 QY 121 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETGLFTLQSELMTVPARGGDP 180
 DB 121 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETGLFTLQSELMTVPARGGDP 180
 QY 181 RPTESCSPGLPRLPALPTAPTPRVWEPVPLEEVQLVVEDEGCAVAPGCTVTITCEVP 240
 DB 181 RPTESCSPGLPRLPALPTAPTPRVWEPVPLEEVQLVVEDEGCAVAPGCTVTITCEVP 240
 QY 241 AQPSTQIHWKDKGVPLIPSPVLLPEIGTQDQGTYSVCVATHSSHCPQESRAVSISITIE 300
 DB 241 AQPSTQIHWKDKGVPLIPSPVLLPEIGTQDQGTYSVCVATHSSHCPQESRAVSISITIE 300
 QY 301 PCEECPTAGSVGGSGICTLALALGILGGIGTA 332
 DB 301 PCEECPTAGSVGGSGICTLALALGILGGIGTA 332
 RESULT 4
 ID AAE23219
 XX AAE23219 standard; Protein: 404 AA.
 XX AC AAE23219;
 XX DT 27-AUG-2002 (first entry)
 XX DE Human receptor for advanced glycosylation end product (PAGE) protein.
 KW Human, Receptor for advanced glycosylation end product; RAGE; cardiant;
 KW tissue growth, neointimal formation, blood vessel; restenosis; diabetes;
 KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;
 KW transgenic animal; acute thrombotic stroke; venous thrombosis.

XX OS Homo sapiens.
 XX PN WO200230889-A2.
 XX PD 18-APR-2002.
 XX PF 12 OCT 2001; 2001WO-US-01036.
 XX PK 13-OCT-2000; 2000US-0687528.
 XX PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX PI Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;
 XX PK: 2002 4-25-00/45.
 XX N-PSDB; AAU06952.
 CC Inhibiting new tissue growth or neointimal formation in blood vessels
 CC of subject suffering from diabetes, stroke and preventing restenosis,
 CC comprises administering inhibitor of receptor for advance glycation end
 CC product
 XX Disclosure, Page 16; 43pp; English.
 CC The invention relates to a method for inhibiting new tissue growth or
 CC neointimal formation in blood vessels in a subject that has experienced
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic
 CC subject. The method comprises administering an inhibitor of receptor for
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new
 CC tissue growth or neointimal formation in subject's blood vessels and
 CC preventing restenosis in the subject. The method is useful for inhibiting
 CC new tissue growth or neointimal formation in blood vessels in a subject
 CC like non-human animal, a transgenic non-human animal or a human suffering
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable
 CC angina, myocardial infarction, abrupt closure following angioplasty or
 CC stent placement, or thrombosis as a result of peripheral vascular surgery.
 CC The method is also useful for preventing restenosis and for determining
 CC whether a compound inhibits new tissue growth in a blood vessel in a
 CC subject. The present sequence is human receptor for advanced
 CC glycosylation end product (RAGE) protein.
 XX Sequence 404 AA;
 Query Match: 99.8%, Score 1762, DB 23, Length 404;
 Best Local Similarity 99.7%, Pred. No. 2.3e-124;
 Matches 331; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQNITARIGEPVLVKCGAPKKPTQPIENKLTGRTTEANKVLSPOGGGPDVSARVLPNG 60
 DB 23 AQNITARIGEPVLVKCGAPKKPTQPIENKLTGRTTEANKVLSPOGGGPDVSARVLPNG 82
 QY 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVQIPCKPIVDASSELTAGVPNKVG 120
 DB 83 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVQIPCKPIVDASSELTAGVPNKVG 142
 QY 121 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETGLFTLQSELMTVPARGGDP 180
 DB 143 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETGLFTLQSELMTVPARGGDP 202
 QY 181 RPTESCSPGLPRLPALPTAPTPRVWEPVPLEEVQLVVEDEGCAVAPGCTVTITCEVP 240
 DB 203 RPTESCSPGLPRLPALPTAPTPRVWEPVPLEEVQLVVEDEGCAVAPGCTVTITCEVP 262
 QY 241 AQPSTQIHWKDKGVPLIPSPVLLPEIGTQDQGTYSVCVATHSSHCPQESRAVSISITIE 300
 DB 263 AQPSTQIHWKDKGVPLIPSPVLLPEIGTQDQGTYSVCVATHSSHCPQESRAVSISITIE 322
 QY 301 PCEECPTAGSVGGSGICTLALALGILGGIGTA 332
 DB 323 PCEECPTAGSVGGSGICTLALALGILGGIGTA 354

Fri May 30 17:16:35 2003

RESULT 5
AAU77543
ID AAU77543 standard; Protein: 404 AA.
XX AC
XX AAU77543;
XX 05-JUN 2002 (first entry)
XX DE
XX Human receptor for advanced glycosylation end product (RAGE).
XX Receptor for advanced glycation end product; RAGE; receptor;
KW amyloid beta peptide; blood-brain barrier; neurovascular stress;
KW cerebral vasoconstriction suppressor; cerebral blood flow enhancer;
KW cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;
KW Alzheimer's disease; Down's syndrome; head trauma; stroke; human.
XX OS Homo sapiens.
XX PN W0200214519 A1.
XX AC
XX 21 FEB 2002.
XX 14-AUG-2001; 2001WO-0525416.
XX 14-AUG 2000; 2000US-063864B.
XX (UYCO) UNIV COLUMBIA NEW YORK.
XX Stern DM, Schmidt AM, Yan SD, Zlokovic B;
PI WPI: 2002-257610/30.
DE N-PSDH; ABK10856.
XX Ameliorating neurovascular stress and decreasing cerebral
PT vasoconstriction in subject suffering from chronic/acute cerebral
PT amyloid angiopathy, by administering inhibitor of receptor for advanced
PT glycation endproduct
XX Disclosure: Page 16; 68pp; English.
XX The invention describes a method of ameliorating neurovascular stress,
CC and decreasing cerebral vasoconstriction in subject suffering from
CC chronic or acute cerebral amyloid angiopathy, comprising administering
CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).
CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain
CC barrier, thus decreasing cerebral vasoconstriction and increasing
CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
CC subject, decreasing cerebral vasoconstriction in a transgenic non-human
CC animal (preferably, transgenic mouse overexpressing mutant human amyloid
CC beta precursor protein) or a human, suffering from chronic or acute
CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy
CC in a subject, where the neurovascular stress is caused by Alzheimer's
CC disease, aging, Down's syndrome, head trauma or stroke. This is the
CC amino acid sequence of human receptor for advanced glycation end
CC product (RAGE) described in the invention.
XX Sequence 404 AA;
XX
Query Match 99.8%; Score 1762; DB 24; Length 404;
Best Local Similarity 99.7%; Pred. No. 2,4e-124;
Matches 331; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AONTATIGELPLVLCVKGAHKPPORLEWKLNTGRTAMKVLSPGGPMDVSARVLPNG 60
ID 23 AONTATIGELPLVLCVKGAHKPPORLEWKLNTGRTAMKVLSPGGPMDVSARVLPNG 82
QY 61 SLFLPVGIGDEGIFRCQAMNRNCKETKSNRYRVYQIPCKPEIVDSASELTAGVFNKVG 120
DB 83 SLFLPVGIGDEGIFRCQAMNRNCKETKSNRYRVYQIPCKPEIVDSASELTAGVFNKVG 142
QY 121 TCNVEGSPYACITLSWHLDDCKPLVINEKGVSVKPGTIRRHPECTLTQLQSELMTVPARGGDP 180
ID 143 TCNVEGSPYACITLSWHLDDCKPLVINEKGVSVKPGTIRRHPECTLTQLQSELMTVPARGGDP 202
QY 181 RPTESCSFSGPLPHRALRTAPTOPRWEPVPLEEVLVVELEGSAVAAGGTVTTCVVP 240
DB 203 RPTESCSFSGPLPHRALRTAPTOPRWEPVPLEEVLVVELEGSAVAAGGTVTTCVVP 262
QY 241 AQPSPQIHWKKGVPPLPSPVLIIPETIGPQDCTYSQVAHSSHCPOESKAVSISTIE 400
DB 263 AQPSPQIHWKKGVPPLPSPVLIIPETIGPQDCTYSQVAHSSHCPOESKAVSISTIE 422
QY 401 PGEEGPTAGSVGGSLGTALALAGLGGGTA 432
DB 433 PGEEGPTAGSVGGSLGTALALAGLGGGTA 454
XX
RESULT 6
AAW44200
ID AAW44200 standard; Protein: 318 AA.
XX AC
XX AAW44200;
XX 14 MAY-1998 (first entry)
XX Human mature receptor to an advanced glycosylation end product.
XX Human; soluble receptor; advanced glycosylation end product; RAGE;
KW AGE; antibody; vascular permeability; diabetes mellitus.
XX OS Homo sapiens.
XX Key location/Qualifiers
DE MISC-difference 66 /note- "encoded by cct"
XX W09739125-A1.
XX 21-OCT-1997.
XX 11-APR-1997; 97WO-EP01834
XX 16-APR-1996; 96US-063314B
XX (SCHD) SCHERING PATENTE AG.
XX Hollander DA, Morser MJ, Nagashima M;
DE WPI: 1997-558380/51.
XX N-PSDH; AAV12395.
XX Anti-advanced glycosylation end product polypeptide antibody
PT prevents receptor binding and therefore reduces vascular
PT permeability, useful to treat diabetes mellitus
XX Claim 2; Page 42-43; 90pp; English.
XX The present sequence represents a mature human receptor to an advanced
CC glycosylation end product (RAGE) polypeptide. The present invention
CC describes an isolated antibody (Ab), specifically immunoreactive with
CC RAGE. Advanced glycosylation end products (AGE) of proteins are
CC non enzymatically glycosylated proteins, which accumulate in vascular
CC tissue in aging, and at an accelerated rate in individuals with
CC diabetes. The Ab, which prevents the interaction between an AGE and its
CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC vascular disorders, neuropathy, retinopathy, nephropathy, hemodialysis
CC associated amyloidosis or atherosclerosis. The Ab can also be used for
CC the isolation and purification of human RAGE polypeptide.
XX Sequence 318 AA;
XX
Query Match 96.3%; Score 1701; DB 18; Length 318;
Best Local Similarity 100.0%; Pred. No. 6,7e-120;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARTIGETVLTAKKAPKPPQPLEWKLNTPTTEAWKVLSPGAGSPWSSVAVPLNG 60
 DB 1 AONTARTIGETVLTAKKAPKPPQPLEWKLNTPTTEAWKVLSPGAGSPWSSVAVPLNG 60
 QY 61 SLFLPAVGIQDERIFKQAMNPNKTKSNRYRVVYQIPKPEIVDSASELTAGVFNKVG 120
 DB 61 SLFLPAVGIQDERIFKQAMNPNKTKSNRYRVVYQIPKPEIVDSASELTAGVFNKVG 120
 QY 121 TCVSEGSYPAGTILSWHLDGKPLVNEKGVSVKPKQIKRHPETGLFTIQSLMLVTPARGGDP 180
 DB 121 TCVSEGSYPAGTILSWHLDGKPLVNEKGVSVKPKQIKRHPETGLFTIQSLMLVTPARGGDP 180
 QY 181 PPTFSCSFSPGLPHHPALPTAPIQPPWEPVPLEEVQIVVEPEGJAVAPGJTVTLTCEVP 240
 DB 181 PPTFSCSFSPGLPHHPALPTAPIQPPWEPVPLEEVQIVVEPEGJAVAPGJTVTLTCEVP 240
 QY 241 AQSPQTHWMKGVLPPLPSPVLLPEIGPQAGGIYSVATHSSHGPPVESKAVSISIIIE 300
 DB 241 AQSPQTHWMKGVLPPLPSPVLLPEIGPQAGGIYSVATHSSHGPPVESKAVSISIIIE 300
 QY 301 PGEESPTAGSVGGSGIGT 318
 DB 301 PGEESPTAGSVGGSGIGT 318

RESULT 7

AAW33754
 ID AAW33754 standard; Protein: 318 AA.

XX AAW33754;

AC 08-MAY-1998 (first entry)

DE Human RAGE polypeptide (318 amino acid residues).

XX Advanced glycosylation end product receptor, RAGE, screening; AGE;
 KW vascular permeability, diabetes mellitus, treatment; atherosclerosis;
 KW Alzheimer's disease.

OS Homo sapiens.

XX W09739121-A1.

XX 23 OCT-1997.

XX 11-APR-1997; 97WO-EP01832.

XX 16-APR-1996; 96US-0633147.

XX (SCHD) SCHERING AG.

XX Morser MJ, Nagashima M;

XX WPI; 1997-526458/48.

XX N-PSDB; AAV06518.

XX New soluble advanced glycosylation end-product receptor polypeptide
 PT - used for reducing vascular permeability, complications of diabetes
 PT etc - also for purification and to screen for modulators

XX Claim 3; Fig 1B; 91pp; English.

XX This is a human advanced glycosylation end product receptor (RAGE)
 CC polypeptide (318 amino acid residues). The RAGE polypeptides and its
 CC active fragments or their mimetics, inhibit interaction between advanced
 CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They
 CC are used to treat diseases associated with AGE/RAGE interaction, such as
 CC increased vascular permeability, diabetes mellitus (particularly
 CC complications such as micro- or macro-vasculopathy or occlusive vascular
 CC disorders such as neuropathy, nephropathy, retinopathy or
 CC atherosclerosis) or hemodialysis-associated amyloidosis, also activation
 CC of microglial cells by beta-amyloid peptides in Alzheimer's disease or

CC age-related disorders such as oxidative stress. These RAGE polypeptides
 CC are also used, when immobilised, to purify AGE from a protein mixture and
 CC to screen for compounds that are agonists and antagonists of AGE/RAGE
 CC interaction. They can also be used diagnostically to detect abnormal
 CC levels of AGE. Antibodies against RAGE polypeptides are useful as
 CC immunoassay reagents for measurement of RAGE levels, and as inhibitors of
 CC interaction between AGE and RAGE or other receptors and for purification
 CC and quantitation of RAGE polypeptides. The encoding nucleic acids are
 CC used to express recombinant RAGE and as probes for isolating related
 CC genes.

XX Sequence 318 AA;

Query Match 95.3%; Score 1781, DB 18, Length 318;

Best local similarity 100.0%; Pred. No. 6, 7e-120;

Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARTIGETVLTAKKAPKPPQPLEWKLNTPTTEAWKVLSPGAGSPWSSVAVPLNG 60

DB 1 AONTARTIGETVLTAKKAPKPPQPLEWKLNTPTTEAWKVLSPGAGSPWSSVAVPLNG 60

QY 61 SLFLPAVGIQDERIFKQAMNPNKTKSNRYRVVYQIPKPEIVDSASELTAGVFNKVG 120

DB 61 SLFLPAVGIQDERIFKQAMNPNKTKSNRYRVVYQIPKPEIVDSASELTAGVFNKVG 120

QY 121 TCVSEGSYPAGTILSWHLDGKPLVNEKGVSVKPKQIKRHPETGLFTIQSLMLVTPARGGDP 180

DB 121 TCVSEGSYPAGTILSWHLDGKPLVNEKGVSVKPKQIKRHPETGLFTIQSLMLVTPARGGDP 180

QY 181 PPTFSCSFSPGLPHHPALPTAPIQPPWEPVPLEEVQIVVEPEGJAVAPGJTVTLTCEVP 240

DB 181 PPTFSCSFSPGLPHHPALPTAPIQPPWEPVPLEEVQIVVEPEGJAVAPGJTVTLTCEVP 240

QY 241 AQSPQTHWMKGVLPPLPSPVLLPEIGPQAGGIYSVATHSSHGPPVESKAVSISIIIE 300

DB 241 AQSPQTHWMKGVLPPLPSPVLLPEIGPQAGGIYSVATHSSHGPPVESKAVSISIIIE 300

QY 301 PGEESPTAGSVGGSGIGT 318

DB 301 PGEESPTAGSVGGSGIGT 318

RESULT 8

AAW44199
 ID AAW44199 standard; Protein: 340 AA.

XX AAW44199;

AC 14-MAY-1998 (first entry)

DE Human soluble receptor to an advanced glycosylation end product.

XX Human, soluble receptor, advanced glycosylation end product, RAGE;
 KW AGE; antibody; vascular permeability; diabetes mellitus.

XX Homo sapiens.

XX W09739125-A1.

XX 23-OCT-1997.

XX 11-APR-1997; 97WO-EP01834.

XX 16-APR-1996; 96US-0633148.

XX (SCHD) SCHERING PATENTE AG.

XX Hollander DA, Morser MJ, Nagashima M;

XX WPI; 1997-558580/51.

XX N-PSDB; AAV12394.

XX Anti-advanced glycosylation end product polypeptide antibody

DE Human SPAGE protein SEQ ID NO 2.
 XX Human: RAGE; receptor for advanced glycosylated endproduct, receptor;
 KW antidiabetic; neuroprotective, cytostatic; antiinflammatory; vasotropic;
 KW nephropathy; dermatologic; antiatherosclerotic; neuropathic; diabetes;
 KW Alzheimer's disease; cancer; inflammation; kidney failure;
 KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.
 OS
 XX Homo sapiens.
 XX WO2001:92892-A2.
 PN 06-DEC-2001
 XX 30-MAY-2001; 2001WO-US17447.
 XX 30-MAY-2000; 2000US-207342P.
 PR 05-MAY-2001; 2001US-0709152.
 XX (TRAN-) TRANS TECH PHARMA.
 PA Shahbaz M;
 PI WPI; 2002-114372/15.
 DR Detecting a receptor for advanced glycosylated endproducts (RAGE)
 PT modulators, for treating e.g., cancer, diabetes or inflammation,
 PT comprises measuring the amount of bound anti-RAGE antibody -
 XX Claim 2; Fig 2; 49pp; English.
 XX The invention relates to detecting receptor for advanced glycosylated
 CC endproducts (RAGE) modulators comprising determining the amount of RAGE
 CC protein or its fragment bound to the pre-adsorbed ligand by measuring the
 CC amount of anti-RAGE antibody bound to the solid surface. The method is
 CC useful for rapid, high-throughput identification of compounds that
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis
 CC or inflammatory lupus nephritis, erectile dysfunction and
 CC atherosclerosis.
 XX Sequence 339 AA;
 SQ Query Match 96.0%; Score 1696; DB 23; Length 339;
 Best Local Similarity 100.0%; Pred No 1,7e-119;
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEKVKLSPOGGPMDSVARVLPNG 60
 DB 23 AQNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEKVKLSPOGGPMDSVARVLPNG 82
 QY 61 SLELPVAVGTDGEGIFRCQAMNNGKETSRYRVYQIPGKPEIVDSASELTAGVPNKVG 120
 DB 83 SLELPVAVGTDGEGIFRCQAMNNGKETSRYRVYQIPGKPEIVDSASELTAGVPNKVG 142
 QY 121 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETCLFTIQLSELMTVPFARGSDP 180
 DB 143 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETCLFTIQLSELMTVPFARGSDP 202
 QY 181 RPEFSCSFSKGLPRHAKLRTAPIQRVWPEVPFLEVLVPEEGAVAPGTVTLTCEVP 240
 DB 203 RPEFSCSFSKGLPRHAKLRTAPIQRVWPEVPFLEVLVPEEGAVAPGTVTLTCEVP 262
 QY 241 AQPSPQTHMKFGVPLPLPPLSHVLILPEIGPDQGTYSVCVATHSHSGPQESRAVSISIE 300
 DB 263 AQPSPQTHMKFGVPLPLPPLSHVLILPEIGPDQGTYSVCVATHSHSGPQESRAVSISIE 322
 QY 301 PGREGPTAGSGGSGIG 317
 DB 323 PGREGPTAGSGGSGIG 339

RESULT 11
 AAE23218
 ID AAE23218 standard: Protein, 416 AA
 XX AAE23218;
 AC AAE23218;
 DT 27-AUG-2002 (first entry)
 XX Cow receptor for advanced glycosylation end product (RAGE) protein.
 XX Cow: Receptor for advanced glycosylation end product; RAGE; cardiant;
 KW +tissue growth; neointimal formation; blood vessel, restenosis; diabetes;
 KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;
 KW transgenic animal; acute thrombotic stroke, venous thrombosis.
 XX Ros taurus.
 OS WO200230889-A2.
 PN 18-APR-2002.
 XX 12-OCT-2001; 2001WO-US32336.
 XX 13-OCT-2000; 2000US-0687328.
 PR (UYCO) UNIV COLUMBIA NEW YORK.
 PA Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;
 PI WPI; 2002-426260/45.
 DR N-PSDB; AAD36951.
 XX Inhibiting new tissue growth or neointimal formation in blood vessels
 CC of subject suffering from diabetes, stroke and preventing restenosis,
 CC comprises administering inhibitor of receptor for advance glycation end
 CC product -
 XX Disclosure; Page 14; 43pp; English.
 XX The invention relates to a method for inhibiting new tissue growth or
 CC neointimal formation in blood vessels in a subject that has experienced
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic
 CC subject. The method comprises administering an inhibitor of receptor for
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new
 CC tissue growth or neointimal formation in subject's blood vessels and
 CC preventing restenosis in the subject. The method is useful for inhibiting
 CC new tissue growth or neointimal formation in blood vessels in a subject
 CC like non-human animal, a transgenic non-human animal or a human suffering
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable
 CC angina, myocardial infarction, abrupt closure following angioplasty or
 CC stent placement, or thrombosis as a result of peripheral vascular surgery
 CC The method is also useful for preventing restenosis and for determining
 CC whether a compound inhibits new tissue growth in a blood vessel in a
 CC subject. The present sequence is cow receptor for advanced glycosylation
 CC end product (RAGE) protein.
 XX Sequence 416 AA;
 SQ Query Match 82.6%; Score 1463; DB 23; Length 416;
 Best Local Similarity 81.5%; Pred No 7.2e-102;
 Matches 278; Conservative 21; Mismatches 30; Indels 12; Gaps 2;
 QY 2 QNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEKVKLSPOGGPMDSVARVLPNGS 61
 DB 24 QNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEKVKLSPOGGPMDSVARVLPNGS 82
 QY 62 LELPVGITDGEIFRCQAMNNGKETSRYRVYQIPGKPEIVDSASELTAGVPNKVG 121
 DB 83 LELPVGITDGEIFRCQAMNNGKETSRYRVYQIPGKPEIVDSASELTAGVPNKVG 142
 QY 122 CVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETCLFTIQLSELMTVPFARGSDP 181
 DB 143 CVSEGSYPAGTILSWHLDGKPLVPNEKGVSKQETRRHPETCLFTIQLSELMTVPFARGSDP 202

CC tissue growth or neointimal formation in subject's blood vessels and
 CC preventing restenosis in the subject. The method is useful for inhibiting
 CC new tissue growth or neointimal formation in blood vessels in a subject
 CC like non-human animal, a transgenic non-human animal or a human suffering
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable
 CC angina, myocardial infarction, abrupt closure following angioplasty or
 CC stent placement, or thrombosis as a result of peripheral vascular surgery
 CC The method is also useful for preventing restenosis and for determining
 CC whether a compound inhibits new tissue growth in a blood vessel in a
 CC subject. The present sequence is mouse receptor for advanced
 CC glycosylation end product (RAGE) protein.

XX SQ Sequence 403 AA;

Query Match 78.3%; Score 1383.5; DB 23; Length 403;
 Best Local Similarity 79.6%; Pred. No. 6.5e-96;
 Matches 262; Conservative 19; Mismatches 47; Indels 1; Gaps 1;
 QY 2 QNITARIGEPVLVKCKGAPKPPQLEWKLNTGRTEAMKVLSPQGGPWDSVAPVLPNGS 61
 Db 24 QNITARIGEPVLVLSCKGAPKPPQLEWKLNTGRTEAMKVLSPQ-GGPWDSVAQILPNGS 82
 QY 62 LFLPVGIDERIFPVLAMNPNKTKSNYPVRYQIPKPEIVDSASELTAGVFNKVG 121
 Db 83 LLLPATGIVDEGTFRCRATNRCKKVKSNYPVRYQIPKPEIVDPASELTASVFNKVG 142
 QY 122 CVSEGSYPAGTILSWHLDGKLLIPDGKETLVKKEETRRHPETGLFTLPSLTVIPTQGGTTH 181
 Db 143 CVSEGSYPAGTILSWHLDGKLLIPDGKETLVKKEETRRHPETGLFTLPSLTVIPTQGGTTH 202
 QY 182 PTFSCSPGLPDRHRLTAPITQPRVPEVPLVEEQVLEVEEGCAVAPCCVTLTCVPA 241
 Db 203 PTFSCSPGLPDRHRLTAPITQPRVPEVPLVEEQVLEVEEGCAVAPCCVTLTCVPA 262
 QY 242 QFSQVHMKKGVEPLPSPVLLIPGEGDQGTYSVAVHSHSGPQESRAVSISIEP 301
 Db 263 QPPQVHMKKGVEPLPSPVLLIPGEGDQGTYSVAVHSHSGPQESRAVSISIEP 322
 QY 302 GEGGTAGSAGSGGCTLALALGILGLG 330
 Db 323 GDEGPAEGSVGESGCTLALALGILGLG 351

RESULT 14

AAU77544
 ID AAU77544 standard; Protein: 403 AA.

AAU77544;

05-JUN-2002 (first entry)

DE Murine receptor for advanced glycosylation end product (RAGE).

XX Receptor for advanced glycation end product; RAGE; receptor;
 KW amyloid beta peptide; blood-brain barrier; neurovascular stress;
 KW cerebral vasoconstriction suppressor; cerebral blood flow enhancer;
 KW cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;
 KW Alzheimer's disease; Down's syndrome; head trauma; stroke; mouse.

OS Mus musculus.

XX AAU77544

WO200214519-A1.

PD 21-FEB-2002.

XX 14-AUG-2001; 2001WO US25416.

XX 14-AUG-2000; 2000US-0638648.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Stern DM, Schmidt AM, Yan SD, Zlokovic H;

DR WPI: 2002-257610/30.
 DR N-PSDB; ABK10857, ABK10858.

XX Ameliorating neurovascular stress and decreasing cerebral
 PT vasoconstriction in subject suffering from chronic/acute cerebral
 PT amyloid angiopathy, by administering inhibitor of receptor for advanced
 PT glycation endproduct

XX Disclosure; Page 17-18; 68pp; English.

XX The invention describes a method of ameliorating neurovascular stress,
 CC and decreasing cerebral vasoconstriction in subject suffering from
 CC chronic or acute cerebral amyloid angiopathy, comprising administering
 CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).
 CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain
 CC barrier, thus decreasing cerebral vasoconstriction and increasing
 CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a
 CC subject, decreasing cerebral vasoconstriction in a transgenic non-human
 CC animal (preferably, transgenic mouse overexpressing mutant human amyloid
 CC beta precursor protein) or a human, suffering from chronic or acute
 CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for
 CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy
 CC in a subject, where the neurovascular stress is caused by Alzheimer's
 CC disease, aging, Down's syndrome, head trauma or stroke. This is the
 CC amino acid sequence of murine receptor for advanced glycation end
 CC product (RAGE) described in the invention.

XX SQ Sequence 403 AA;

Query Match 78.3%; Score 1383.5; DB 23; Length 403;
 Best Local Similarity 79.6%; Pred. No. 6.5e-96;
 Matches 262; Conservative 19; Mismatches 47; Indels 1; Gaps 1;
 QY 2 QNITARIGEPVLVKCKGAPKPPQLEWKLNTGRTEAMKVLSPQGGPWDSVAPVLPNGS 61
 Db 24 QNITARIGEPVLVLSCKGAPKPPQLEWKLNTGRTEAMKVLSPQ-GGPWDSVAQILPNGS 82
 QY 62 LFLPVGIDERIFPVLAMNPNKTKSNYPVRYQIPKPEIVDSASELTAGVFNKVG 121
 Db 83 LLLPATGIVDEGTFRCRATNRCKKVKSNYPVRYQIPKPEIVDPASELTASVFNKVG 142
 QY 122 CVSEGSYPAGTILSWHLDGKLLIPDGKETLVKKEETRRHPETGLFTLPSLTVIPTQGGTTH 181
 Db 143 CVSEGSYPAGTILSWHLDGKLLIPDGKETLVKKEETRRHPETGLFTLPSLTVIPTQGGTTH 202
 QY 182 PTFSCSPGLPDRHRLTAPITQPRVPEVPLVEEQVLEVEEGCAVAPCCVTLTCVPA 241
 Db 203 PTFSCSPGLPDRHRLTAPITQPRVPEVPLVEEQVLEVEEGCAVAPCCVTLTCVPA 262
 QY 242 QFSQVHMKKGVEPLPSPVLLIPGEGDQGTYSVAVHSHSGPQESRAVSISIEP 301
 Db 263 QPPQVHMKKGVEPLPSPVLLIPGEGDQGTYSVAVHSHSGPQESRAVSISIEP 322
 QY 302 GEGGTAGSAGSGGCTLALALGILGLG 330
 Db 323 GDEGPAEGSVGESGCTLALALGILGLG 351

RESULT 15

AAU48747
 ID AAU48747 standard; protein; 112 AA.

AAU48747;

02-APP-2002 (first entry)

XX Human RAGE protein V-domain SEQ ID NO 3.

XX Human: RAGE; receptor for advanced glycated endproduct; receptor;
 KW antidiabetic; neuroprotective; cytostatic; antiinflammatory; vasotropic;
 KW nephrotropic; dermatological; antiarteriosclerotic; nontropic; diabetes;
 KW Alzheimer's disease; cancer; inflammation; kidney failure;
 KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.

Core version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:55:12, Search time 26.9189 seconds
(without alignments)
362.883 Million cell updates/sec

Title: US-09-872-185b-2
Perfect score: 1766
Sequence: 1 AONTAR:GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG 332

Scoring table: BLOSUM62
Gap: 10 0 0, Gap: 10 0 0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*
3: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*
4: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*
5: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*
6: /cgn2_6/ptedata/1/iaa/5A_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	332	US-08-062-365-1	Sequence 1, Appli
2	1701	96.3	318	US-08-633-148-4	Sequence 4, Appli
3	1701	96.3	340	US-08-633-148-2	Sequence 2, Appli
4	1495	84.7	278	US-08-432-016-5	Sequence 5, Appli
5	1495	84.7	278	US-08-684-594-5	Sequence 5, Appli
6	218.5	12.4	583	US-08-432-016-2	Sequence 2, Appli
7	218.5	12.4	583	US-08-684-594-2	Sequence 2, Appli
8	212.5	12.0	456	US-08-432-016-4	Sequence 4, Appli
9	212.5	12.0	456	US-08-684-594-4	Sequence 4, Appli
10	209	11.8	1651	US-09-540-245A-18	Sequence 18, Appli
11	205	11.6	477	US-08-432-016-3	Sequence 3, Appli
12	205	11.6	477	US-08-684-594-3	Sequence 3, Appli
13	197	11.2	486	US-08-432-016-6	Sequence 6, Appli
14	197	11.2	486	US-08-684-594-6	Sequence 6, Appli
15	190.5	10.8	698	US-08-602-725-16	Sequence 36, Appli
16	190.5	10.8	734	US-08-389-459A-17	Sequence 17, Appli
17	190.5	10.8	744	US-08-987-867A-17	Sequence 17, Appli
18	189.5	10.7	1297	US-09-540-245A-17	Sequence 17, Appli
19	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
20	177	10.0	924	US-08-656-984A-28	Sequence 28, Appli
21	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
22	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
23	176.5	10.0	421	US-08-659-984A-1	Sequence 1, Appli
24	176.5	10.0	421	US-08-660-531-1	Sequence 1, Appli
25	176.5	10.0	444	US-08-659-984A-5	Sequence 5, Appli
26	176.5	10.0	444	US-08-660-531-5	Sequence 5, Appli
27	172.5	9.8	642	US-08-217-299-1	Sequence 1, Appli

28	172	9.7	1266	4	US-08-506-296B-4	Sequence 4, Appli
29	170	9.6	462	2	US-08-752-307B-7	Sequence 7, Appli
30	170	9.6	462	4	US-09-707-802-7	Sequence 7, Appli
31	170	9.6	462	4	US-09-941-426-7	Sequence 7, Appli
32	170	9.6	465	2	US-08-752-307B-5	Sequence 5, Appli
33	170	9.6	465	4	US-09-707-802-5	Sequence 5, Appli
34	170	9.6	465	4	US-09-941-426-5	Sequence 5, Appli
35	168	9.5	1005	4	US-09-510-245A-15	Sequence 15, Appli
36	165	9.3	607	2	US-08-752-307B-12	Sequence 12, Appli
37	165	9.3	607	4	US-09-707-802-12	Sequence 12, Appli
38	165	9.3	607	4	US-09-941-426-12	Sequence 12, Appli
39	164	9.3	1091	4	US-08-486-485-5	Sequence 5, Appli
40	164	9.3	1447	4	US-09-041-886-25	Sequence 25, Appli
41	164	9.3	1447	5	US-09-041-886-25	Sequence 25, Appli
42	162	9.2	30	4	US-09-062-365-5	Sequence 5, Appli
43	159.5	9.0	1101	3	US-08-586-485-2	Sequence 2, Appli
44	158.5	9.0	1253	4	US-08-506-296B-14	Sequence 14, Appli
45	158	8.9	612	2	US-08-752-307B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-09-062-365-1
: Sequence 1, Application US/J9062365
: Patent No. 6465422
: GENERAL INFORMATION:
: APPLICANT: Schmidt, Ann Marie
: TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A
: TITLE OF INVENTION: SUBJECT
: REFERENCE: 55124
: CURRENT APPLICATION NUMBER: US/09/062,365
: CURRENT FILING DATE: 1998-04-17
: NUMBER OF SEQ. IDS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ. ID NO. 1
: LENGTH: 332
: TYPE: PRT
: ORGANISM: Human
US-09-062-365-1

Query Match	100.0%	Score	1766	DB	4	Length	332
Best Local Similarity	100.0%	Pred. No.	1.2e-141				
Matches	332	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Q1	1	AONTAR:GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG	60				
D1	1	AONTAR:GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG	60				
Q2	61	SLFLPVGIGDEIFRCQAMNRKETSINRYRVYQIFGKPEIVDSASELTAGVKNVG	120				
D2	61	SLFLPVGIGDEIFRCQAMNRKETSINRYRVYQIFGKPEIVDSASELTAGVKNVG	120				
Q3	121	TCVSEGSYPAGTISLWDLGKPLVPNEKYSVKRPHRPHETGIFLLOSHLWTFAGGUP	180				
D3	121	TCVSEGSYPAGTISLWDLGKPLVPNEKYSVKRPHRPHETGIFLLOSHLWTFAGGUP	180				
Q4	181	RPTFSQSPSGLPPHFLPTAPIQPPVWEPVLEFVQIVVEPCCAVAPGCTVITCEVP	240				
D4	181	RPTFSQSPSGLPPHFLPTAPIQPPVWEPVLEFVQIVVEPCCAVAPGCTVITCEVP	240				
Q5	241	AQSTFQIIMWVIFVVFPLFSSPVLLIPELFGVQNTYSVATISSHPTSPSPAVSISITIE	400				
D5	241	AQSTFQIIMWVIFVVFPLFSSPVLLIPELFGVQNTYSVATISSHPTSPSPAVSISITIE	400				
Q6	301	FGFEGTADSSVSSG:GILALAGTICGLTIA	332				
D6	301	FGFEGTADSSVSSG:GILALAGTICGLTIA	332				
Q7	301	FGFEGTADSSVSSG:GILALAGTICGLTIA	332				
D7	301	FGFEGTADSSVSSG:GILALAGTICGLTIA	332				

RESULT 2

```

US 08 644 148 4
: Sequence 4, Application US/08644148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSE, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS Dos
: SOFTWARE: Patent In Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08644148
: FILING DATE: 16 APR 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY ESQ., MATTHEW B.
: REGISTRATION NUMBER: 39,787
: REFERENCE/DOCKET NUMBER: 014618-00560005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 418 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US 08-644-148 4

```

```

Query Match 96.48; Score 1701; DB 2; Length 418;
Best local Similarity 100.0%; Pred. No. 3,76-146;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 60
DB 1 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 60

QY 61 SLFLPVGTLQDEGIFRQVQAMNNPKETKSNRYRVVYQIPGKPEIVDSASELTAGVPNKVG 120
DB 61 SLFLPVGTLQDEGIFRQVQAMNNPKETKSNRYRVVYQIPGKPEIVDSASELTAGVPNKVG 120

QY 121 TCVSQSGYPACILSWHLKPKPLVNEKGVSKQTRRHPTGLFTLQSELMVTIPARGGDP 180
DB 121 TCVSQSGYPACILSWHLKPKPLVNEKGVSKQTRRHPTGLFTLQSELMVTIPARGGDP 180

QY 181 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 240
DB 181 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 240

QY 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPDQKTYSCVATHSSHPQESRAVSTSLIE 400
DB 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPDQKTYSCVATHSSHPQESRAVSTSLIE 400

QY 401 PGEGPTAGSVGGSGIGT 418
DB 401 PGEGPTAGSVGGSGIGT 418

```

```

RESULT 4
US-08-644-148 2
: Sequence 4, Application US/08644148

```

```

: Sequence 2, Application US/08644148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSE, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS Dos
: SOFTWARE: Patent In Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08644148
: FILING DATE: 16 APR 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY ESQ., MATTHEW B.
: REGISTRATION NUMBER: 39,787
: REFERENCE/DOCKET NUMBER: 014618-00560005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US 08-644-148-2

```

```

Query Match 96.48; Score 1701; DB 2; Length 340;
Best local Similarity 100.0%; Pred. No. 46-146;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 60
DB 1 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 60

QY 23 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 62
DB 23 AONTTARIGEFVLKKGAKKPKPQLKWKLTGRTKAWKVLSPGSGGWDSVARVLPNG 62

QY 61 SLFLPVGTLQDEGIFRQVQAMNNPKETKSNRYRVVYQIPGKPEIVDSASELTAGVPNKVG 120
DB 61 SLFLPVGTLQDEGIFRQVQAMNNPKETKSNRYRVVYQIPGKPEIVDSASELTAGVPNKVG 120

QY 121 TCVSQSGYPACILSWHLKPKPLVNEKGVSKQTRRHPTGLFTLQSELMVTIPARGGDP 180
DB 121 TCVSQSGYPACILSWHLKPKPLVNEKGVSKQTRRHPTGLFTLQSELMVTIPARGGDP 180

QY 181 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 240
DB 181 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 240

QY 203 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 262
DB 203 RPTFSCSFSPGLPRHRLATAPIQPRVWEPVPLEEVQIVVEPRGAVAPGTVITICEVP 262

QY 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPDQKTYSCVATHSSHPQESRAVSTSLIE 400
DB 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPDQKTYSCVATHSSHPQESRAVSTSLIE 400

QY 401 PGEGPTAGSVGGSGIGT 418
DB 401 PGEGPTAGSVGGSGIGT 418

```

```

RESULT 4
US 08 644-016-5
: Sequence 5, Application US/08644148

```

Patent No. 5968768
 GENERAL INFORMATION:
 APPLICANT: HAYNES, BARTON F.
 APPLICANT: AROFFO, ALEJANDRO
 APPLICANT: PATEL, DHAVALKUMAR
 APPLICANT: BOWEN, MICHAEL A.
 APPLICANT: MARQUARDT, HANS
 TITLE OF INVENTION: CD6 LIGAND
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHYE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/432,016
 FILING DATE: 01-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/333,350
 FILING DATE: 02-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/143,903
 FILING DATE: 02-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: WILSON, MARY J.
 REGISTRATION NUMBER: 32,955
 REFERENCE/DOCKET NUMBER: 1579-95
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 278 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-432-016-5

Query Match 84.7%; Score 1495; DB 2; Length 278;
 Best Local Similarity 99.6%; Pred. No. 8e-119;
 Matches 277; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 8 IGEPLVLKCKGAPKPPQRLKWKLNICRTTEAKVKVLSPOGGPWSVAVPLPNCGLFLPAV 67
 DB 1 IGEPLVLKCKGAPKPPQRLKWKLNICRTTEAKVKVLSPOGGPWSVAVPLPNCGLFLPAV 60
 QY 68 GIGDGGIFPCQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 127
 DB 61 GIGDGGIFPCQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 120
 QY 128 YPAGTLSWHLIDGKPLVFNKGVSVKQETRRHPETGLFTIQSELMTVPARGSDPRPTFSCS 187
 DB 121 YPAGTLSWHLIDGKPLVFNKGVSVKQETRRHPETGLFTIQSELMTVPARGSDPRPTFSCS 180
 QY 188 FSPGLPFRHRLTAPIQPRVWEPVPLEEVQLVVEPEGGAAPGGIVLTICEVPAQPSQI 247
 DB 181 FSPGLPFRHRLTAPIQPRVWEPVPLEEVQLVVEPEGGAAPGGIVLTICEVPAQPSQI 240
 QY 248 HMKDGVPLPLPSPVLIIFPIGPGQGGTYSVCVATHSS 285
 DB 241 HMKDGVPLPLPSPVLIIFPIGPGQGGTYSVCVATHSS 278

RESULT 5

US-08-684-594-5
 Sequence 5, Application US/08684594
 Patent No. 5998172
 GENERAL INFORMATION:
 APPLICANT: HAYNES, BARTON F.
 APPLICANT: AROFFO, ALEJANDRO
 APPLICANT: PATEL, DHAVALKUMAR
 APPLICANT: BOWEN, MICHAEL A.
 APPLICANT: MARQUARDT, HANS
 TITLE OF INVENTION: CD6 LIGAND
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHYE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/684,594
 FILING DATE: 18-JUL-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/432,016
 FILING DATE: 01-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/333,350
 FILING DATE: 02-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/143,903
 FILING DATE: 02-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: WILSON, MARY J.
 REGISTRATION NUMBER: 32,955
 REFERENCE/DOCKET NUMBER: 1579-112
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 278 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-684-594-5

Query Match 84.7%; Score 1495; DB 2; Length 278;
 Best Local Similarity 99.6%; Pred. No. 8e-119;
 Matches 277; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 8 IGEPLVLKCKGAPKPPQRLKWKLNICRTTEAKVKVLSPOGGPWSVAVPLPNCGLFLPAV 67
 DB 1 IGEPLVLKCKGAPKPPQRLKWKLNICRTTEAKVKVLSPOGGPWSVAVPLPNCGLFLPAV 60
 QY 68 GIGDGGIFPCQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 127
 DB 61 GIGDGGIFPCQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 120
 QY 128 YPAGTLSWHLIDGKPLVFNKGVSVKQETRRHPETGLFTIQSELMTVPARGSDPRPTFSCS 187
 DB 121 YPAGTLSWHLIDGKPLVFNKGVSVKQETRRHPETGLFTIQSELMTVPARGSDPRPTFSCS 180
 QY 188 FSPGLPFRHRLTAPIQPRVWEPVPLEEVQLVVEPEGGAAPGGIVLTICEVPAQPSQI 247
 DB 181 FSPGLPFRHRLTAPIQPRVWEPVPLEEVQLVVEPEGGAAPGGIVLTICEVPAQPSQI 240
 QY 248 HMKDGVPLPLPSPVLIIFPIGPGQGGTYSVCVATHSS 285

31 VNSAYGDTIIIPCP---LDVQNLMF---GK---WKYEKPSPEVIAFFSTKKSVMY 80
Db
51 DSV AKVLINGSLFPAVCIQUCIQRGAMNKNCKTKSNY----RVRYQIIP 99
QY
81 DDVPEYKORLNSENYTLSTISNARISDEKFEVCLV-----TEDNVFEATPIVKVEKP 134
Db
100 KRPFTVTSASLILAVENKVIQVWSEYSGFAGTISWHLEKPLVPNKEKVSRYKREATPRIP 159
QY
135 SKPEIVSKALELEINQKKKQDGLSELSYFQGNILWYKNGKVIHPLLEGAAVILFKKEMIP 194
Db
160 EPGLETLASLMTVPARGGRRPPTFSCS---FSPGLPPHPALDAPIQPVME-PVPLEE 215
QY
195 VTQLTMTSTLEYKITK-ADIQMPTCSVTYAPS-----GQKTHISEQAVDEIYYPTEQ 248
Db
216 VOLVVERECAAVAGCVITLQVTAQIST-QIHWKKGVLPLPFSVLILFELGPDQC 274
QY
249 VTIVLPPKXAIKEQDNIILKLGKNRPPEEFELVLPQPEEIPFSNTYTLIXDVFNAI 308
Db
275 RYSG 279
QY
309 GDYKC 313
Db

RESULT 8

US-08-432-016-4

```

: Sequence 4. Application US/05442015
: Patent No. 5968768
: GENERAL INFORMATION
: APPLICANT: HAYNES, BARTON F.
: APPLICANT: ARUFFO, ALEJANDRO
: APPLICANT: PATEL, DHAAVALKUMAR
: APPLICANT: BOWEN, MICHAEL A.
: APPLICANT: MARQUARDT, HANS
: TITLE OF INVENTION: CD6 LIGAND
: NUMBER OF SEQUENCES: 9
: CORRESPONDENT'S ADDRESS
: ADDRESSEE: NIXON & VANDERHYE P. C.
: STREET: 1100 NORTH GLEBE ROAD
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: U.S.A.
: ZIP: 22201-4714
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0.
: CURRENT APPLICATION DATA: US/08/432,016
: FILING DATE: 01-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/333,350
: FILING DATE: 02 NOV 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/143,903
: FILING DATE: 02 NOV 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: WILSON, MARY J.
: REGISTRATION NUMBER: 32,955
: REFERENCE/DOCKET NUMBER: 1579-95
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 466 amino acids
: TYPE: amino acid
: STRANDNESS: linear
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: PS-08-432-016-4

```

```

Query Match      12 Q#; Score 312.5; DB 2; Length 466;
Best local Similarity 25 5#; pred. No. 4.6e-10;
Matches 76; Conservative 40; Mismatches 133; Indels 49; Gaps 11;

          QY    9 GEPLVKCKGAPKPPORLEWKLTGRTEAWKVLSGGGPMDSVAR----- 55
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          RB    2 GETIVVPNDGTKKITDGLITFK-----WKVVK-DDGSPGDLLVKQAQKD EATVSATD 52
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          QY    56 -----VLNWSSTPIFAVGILDEISPELWAMFNPKFRIKSNYPVVVQLPCKRPFLVISA 108
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          Db     53 GKYSPTSVAANSSLLIARGSLADRGVFCTMVVSEFTINLEEYS-VVEVKVRHKPSAPIKRNA 111
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          QY    109 SELTASGVNKVIGICVSEGSYPASTLSWHLDGKELFVENEGSVSKVEUTRHPHTGLTFIQS 168
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          Db     112 KELENGRLTGAGECVENANPADLIWKKNQTILDGKTIITISIIIKDKITGLSSTSS 171
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          QY    149 FLMTTPAQCEFFPTSCSESPILPRIHALPTAFIPFRWEPE-----VFLEEVLVVEEP 222
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          Db     172 KLYNT ARKEDEVESCFCTI-----AKH---VMGPdq--VSEFSEFIHYPTKESLQVVS 220
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          QY    223 EGCAVAGGTVTITCEVPAQHSP-QIRHWMDKGVLPLPHPSVPIILDFEIGPDQGTYSC 279
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          Db     221 Q'-SPIREGEDVTLKQADGNPPPTSPFNILKGGKVTVTDQDVYTITGVTRADSGIYKC 277
              ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

RESULT 9

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


QY 179 DRRPTFSCSPGLPHEALPTAPICQVWVE-PVPLEEVQVLLVPEEGCAVALGCTVTLTC 237
 Db 179 DANAKFTCIYVHGPSQ--KTUSEEVFVHVPTKVTIRVLSUSSTIKEDGNVTLKC 236
 QY 238 EVPAQSPQ----- 246
 Db 237 SCGNPPPPQFFLYIPGTEGIRSSQTYVYVMTDVPNNATGEYKCSLLDKSMDDTTITVHY 296
 QY 247 -----IHWKKGVPPLPLPSPVLLPEIGPQDQ 274
 Db 237 LDLOITRSGEVTKQJGEALFVSQTISSSNAIVFWIKNTKMTSPS---FSSLOYQDA 352
 QY 275 GTYSVATHSHGHQESKAVSISLIE 300
 Db 353 GNYICETTHKEVEGLKKRKLKIVE 378

RESULT 12
 US-08-684-594-3
 ; Sequence 1, Application US/08684594
 ; Patent No. 5998172
 ; GENERAL INFORMATION:
 ; APPLICANT: HAYNES, BARTON F.
 ; APPLICANT: ARUFFO, ALEJANDRO
 ; APPLICANT: PATEL, DHAVALKUMAR
 ; APPLICANT: BOWEN, MICHAEL A.
 ; APPLICANT: MARQUARDT, HANS
 ; TITLE OF INVENTION: CD6 LIGAND
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDEPHYE, P.C.
 ; STREET: 1100 NORTH GLAHE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/684,594
 ; FILING DATE: 18-JUL-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/432,016
 ; FILING DATE: 01-MAY-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/333,350
 ; FILING DATE: 02-NOV-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/143,903
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-112
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 477 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-684-594-3

Query Match 11.6%; Score 205; DB 2; Length 477.
 Best Local Similarity 21.5%; Pred. No. 2e-09;

Matches 70; Conservative 39; Mismatches 125; Indels 92; Gaps 6.
 QY 59 NGSLLFAVILQDEIFRQJAMNKN:KETKSNVFPVYUJPGKPEIVDSASSETLAGVPMK 118
 Db 61 NYLLISINARISDEKREVCMLVTEDDVSEPTV-KVFKQPSQPEILHQADELETEKLM 119
 QY 119 VGTCSGSGSPAGTSLSHLDGKPLVPNEKGVSKQTRRHPTETGLFTLQSELMVTPAKGG 178
 Db 120 LGEVWDSYFEGNVWTKYKNGRVLPQVEVVVNLKRVENRSTGLFTMTSSSLQYMPK-E 178
 QY 179 DPPFTFSCSPGLPHEALPTAPICQVWVE-PVPLEEVQVLLVPEEGCAVALGCTVTLTC 237
 Db 179 DANAKFTCIYVHGPSQ--KTUSEEVFVHVPTKVTIRVLSUSSTIKEDGNVTLKC 236
 QY 238 EVPAQSPQ----- 246
 Db 237 SCGNPPPPQFFLYIPGTEGIRSSQTYVYVMTDVPNNATGEYKCSLLDKSMDDTTITVHY 296
 QY 247 -----IHWKKGVPPLPLPSPVLLPEIGPQDQ 274
 Db 237 LDLOITRSGEVTKQJGEALFVSQTISSSNAIVFWIKNTKMTSPS---FSSLOYQDA 352
 QY 275 GTYSVATHSHGHQESKAVSISLIE 300
 Db 353 GNYICETTHKEVEGLKKRKLKIVE 378

RESULT 13
 US-08-432-016-6
 ; Sequence 6, Application US/08432016
 ; Patent No. 598768
 ; GENERAL INFORMATION:
 ; APPLICANT: HAYNES, BARTON F.
 ; APPLICANT: ARUFFO, ALEJANDRO
 ; APPLICANT: PATEL, DHAVALKUMAR
 ; APPLICANT: BOWEN, MICHAEL A.
 ; APPLICANT: MARQUARDT, HANS
 ; TITLE OF INVENTION: CD6 LIGAND
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDEPHYE, P.C.
 ; STREET: 1100 NORTH GLAHE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/432,016
 ; FILING DATE: 01-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/333,350
 ; FILING DATE: 02-NOV-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/143,903
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WILSON, MARY J.
 ; REGISTRATION NUMBER: 32,955
 ; REFERENCE/DOCKET NUMBER: 1579-95
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4000
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 486 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:

[illegible]


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US-09-872-185b-2
; FILING DATE: 02-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/01816
; FILING DATE: 19-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9317423
; FILING DATE: 21-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36663
; REFERENCE/INVENTOR NUMBER: 1090-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4091
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 698 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-872-185b-2

```

```

Query Match      10.88; Score 190.5; DB 2; Length 698;
Best local Similarity 17.44; Prod No 5.8e-08;
Matches 89; Conservative 62; Mismatches 134; Indels 21; Gaps 16.

QY 3 NITAPICEPLVLCGAPKKPQPPQLEWKLNTGTEAWKVLSPQGGGPPQDSVAPVLPNGSL 62
Db 242 NTSYSEGENLSCHAA-SNFPQYSKFN-----GTGCGSTQ-----EL 280

QY 63 FLAVAGLQDRGIEKQAMNEN-CKEIKSNYRVKVYQIPCKKIVDSASF----- 110
Db 281 FIPNITVNSGSYTCQHNSDTHLNPPTTTTIVAFPPKPTITSNNSNPVEDDAVALI 340

QY 111 -----LI-----AGVENKVGTCVS 124
Db 341 CEPELQNTVYIWWNNGSLIVSPHQLSLNINRTLLLSVTRNIVGPGYEGGIONELSVIHS 400

QY 125 E-----GSYPAGTLSMHLGDKPLVPNEKGV 150
Db 401 DPVILNVLYGPDPTISPSYTYRPGVNLSSLSCHAAASNPAPQYSLIDG-----N 450

QY 151 VKQTRR-----HPTGCTLTQSLMVTTPAGCG-----PRTFSCSPSP 190
Db 451 IQHTQELFISNTERNSGLYTCQAN---NSASGHSRTTVKTIIVSAELPRDSISSNSK 507

QY 191 GLPFRAL-----FTAPIQPPVW----- 208
Db 508 PVKIDAVATQCPHAGNFIYIWWNNGSLIVSPHQLSLNINRTLLLSVTRNIVGPGYEGGIONELSVIHS 557

QY 209 -----EPVPLEEV-----QIVVFPFGAVAPAGSTVILTCEVPAGPSPQIHWMDKG 253
Db 568 GIQNSVANSDDVTLGVLYGPDPTISPPDSYLSGANLSCHASNPSPQYSWRING 627

QY 254 VPLPSPVPLILFEICQDQGGIYSVAIHSSHGPGCFSPKAVSISIEPCEFGPTAGSVGG 313
Db 628 I--PQHTQVLFIAKTIPTNNNGTYACFVSNATGRNNSIVKSTIVSASGT-----SPGL 679

QY 314 SGLGTALALGILGL 329
Db 680 SACATVGMIGVAVGV 695

```

Search completed: May 30, 2003, 15:59:51
Job time : 27.9189 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 30, 2003, 17:16:37, Seq: ch. 12, Seq: ch. 12, 1712 S. reads
(without alignments)
1284.095 Million cell updates/sec

Title: US-09-872-185B 2

Perfect score: 1766

Sequence: 1 AQLNIAKIGHIVLAKGAP NSGSHIAIAIAIAIAIAIAIA 332

Scoring table: BLOSUM62

Gap: 10 0 Gap: 10 5

Searched: 383519 seqs, 101231694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB_PEP*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB_PEP*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB_PEP*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB_PEP*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB_PEP*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB_PEP*
- 7: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB_PEP*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB_PEP*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB_PEP*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB_PEP*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB_PEP*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB_PEP*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB_PEP*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB_PEP*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	332	9 US-09-851-071-1	Sequence 1, Appli
2	1766	100.0	332	10 US-09-872-185B-2	Sequence 2, Appli
3	1751.5	99.2	405	1 US-08-755-235 4	Sequence 4, Appli
4	1463	82.8	416	1 US-08-755-235 2	Sequence 2, Appli
5	595	33.7	112	10 US-09-872-185B 1	Sequence 1, Appli
6	220.5	12.5	582	9 US-09-736-457-334	Sequence 334, App
7	220.5	12.5	582	9 US-09-902-941-331	Sequence 334, App
8	220.5	12.5	582	9 US-09-849-626-334	Sequence 334, App
9	220.5	12.5	582	9 US-10-017-754-334	Sequence 334, App
10	219.5	12.4	583	9 US-10-176-847-64	Sequence 64, Appli
11	201	11.4	793	9 US-09-808-602-70	Sequence 70, Appli
12	201	11.4	793	9 US-09-808-193-23	Sequence 23, Appli
13	201	11.4	793	9 US-09-800-194-54	Sequence 59, Appli
14	192.5	10.9	1252	9 US-09-968-193 19	Sequence 19, Appli
15	192.5	10.9	1253	9 US-09-968-193 20	Sequence 20, Appli
16	190.5	10.8	734	10 US-09-756-551A-17	Sequence 17, Appli
17	189.5	10.7	484	9 US-09-977-033A-24	Sequence 24, Appli
18	189.5	10.7	404	9 US-09-977-033A-24	Sequence 24, Appli
19	189.5	10.7	404	9 US-10-161-572 44	Sequence 44, Appli

ALIGNMENTS

RESULT 1

US-09-851-071-1

: Sequence 1, Application US/98851071

: Patent No. US2001/07550A1

: GENERAL INFORMATION:

: APPLICANT: Schmidt, Anne Marie

: APPLICANT: Stern, David

: TITLE OF INVENTION: A METHOD OF INHIBITING TUMOR INVASION OF SPREADING IN A SUB-E

: FILE REFERENCE: 0575/55424-Z/JPW/SHS/MVM

: CURRENT APPLICATION NUMBER: US/09/851,071

: CURRENT FILING DATE: 2001-05-08

: NUMBER OF SEQ ID NOS: 6

: SOFTWARE: Patent in version 3.1

: SEQ ID NO 1

: LENGTH: 332

: TYPE: PRT

: ORGANISM: Human

: US-09-851-071-1

Query Match: 100.0%; Score 1766; DB %: Length 332;

Best Local Similarity: 100.0%; Pred. No. 8.4e-97;

Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY	1	AQNIAPGPPVILKKGAPKPPKPLEWKLNGITFAWKVSPQGGHWSVAVPLNG	60
DB	1	AQNIAPGPPVILKKGAPKPPKPLEWKLNGITFAWKVSPQGGHWSVAVPLNG	60
QY	61	SLFLPAGVIGDEGIFECQAMNRNGKIKSNFYVYQIPKPEIVDSASELTAGYVNVKG	120
DB	61	SLFLPAGVIGDEGIFECQAMNRNGKIKSNFYVYQIPKPEIVDSASELTAGYVNVKG	120
QY	121	ICVSESGYPASTLSWLDGKPLVPNEKGVSVREYTKRHETGLFTLQSELMTYPAKGDP	180
DB	121	ICVSESGYPASTLSWLDGKPLVPNEKGVSVREYTKRHETGLFTLQSELMTYPAKGDP	180
QY	181	RPTFSGSPCLDRHPRALPTAPIQPVWEPVPLEEVOLVWPEGCAVAGCTVTLCVPT	240
DB	181	RPTFSGSPCLDRHPRALPTAPIQPVWEPVPLEEVOLVWPEGCAVAGCTVTLCVPT	240
QY	241	AQSPFLHMMKGVGLPPLPSPVLLIPPLPPLPPLPPLPPLPPLPPLPPLPPLPPLP	300
DB	241	AQSPFLHMMKGVGLPPLPSPVLLIPPLPPLPPLPPLPPLPPLPPLPPLPPLPPLP	300

```

QY      301 PGEKPTAGSVGSGSLGLALALAGLGCGTA 432
Db      |||||||
Db      301 PGEKPTAGSVGSGSLGLALALAGLGCGTA 432
|||

RESULT 2
US 09-872,185B 2
? Sequence 2, Application US/098/2185B
? Patent No. US2002122799A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Herold, Kevin
? APPLICANT: Yan, Shi Du
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Lamster, Ira
? TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
? FILE REFERENCE: 0575/64080
? CURRENT APPLICATION NUMBER: us 09/872,185B
? CURRENT FILING DATE: 2001-06-01
? NUMBER OF SEQ ID NOS: 16
? SOFTWARE: Patent In Version 4.1
? SEQ ID NO 2
? LENGTH: 432
? TYPE: PRF
? ORGANISM: Human
US 09-872,185B 2

Query Match          100.0%   Score 1746; DB 10; Length 332;
Best Local Similarity 100.0%; Prod. No. 8.4e+97;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AONTTARICEPIVAKTKCATKKPQRGRWKLNIQRIFAWKVSPQGSHDVSAPVLINC 60
Db      |||||||
Db      1 AONTTARICEPIVAKTKCATKKPQRGRWKLNIQRIFAWKVSPQGSHDVSAPVLINC 60
|||||

QY      61 SLELPVAVQLQEGIRIQVAMNPNCKETKSYPVPVVQTCPKPPIVDISAELTAGVFNKV 120
Db      |||||||
Db      61 SLELPVAVQLQEGIRIQVAMNPNCKETKSYPVPVVQTCPKPPIVDISAELTAGVFNKV 120
|||||

QY      121 TCVSRSYSFAGLISWHLLKIKLVINEKCVSKQIIRHPEIGLFIQSRLMVIPARGDDP 180
Db      |||||||
Db      121 TCVSRSYSFAGLISWHLLKIKLVINEKCVSKQIIRHPEIGLFIQSRLMVIPARGDDP 180
|||||

QY      181 RPTSCSFSSRGLPRHRALRIAPQPRVWEIVPLEEQULVVEPKGAVAPGGCIVTLCEVP 240
Db      |||||||
Db      181 RPTSCSFSSRGLPRHRALRIAPQPRVWEIVPLEEQULVVEPKGAVAPGGCIVTLCEVP 240
|||||

QY      241 AGPSFQIHMKGQVPLFILPSPVLLDEICPDQGTYSVVAIHSSHGPOESRAVSISTIE 400
Db      |||||||
Db      241 AGPSFQIHMKGQVPLFILPSPVLLDEICPDQGTYSVVAIHSSHGPOESRAVSISTIE 400
|||||

QY      301 PGEKPTAGSVGSGSLGLALALAGLGCGTA 432
Db      |||||||
Db      301 PGEKPTAGSVGSGSLGLALALAGLGCGTA 432
|||

RESULT 3
US 08 755,245 4
? Sequence 4, Application US/0875/245
? Publication No. US2004005942A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Wu, Jun
? TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
? FILE REFERENCE: 0575/50194
? CURRENT APPLICATION NUMBER: US 08/755,245
? CURRENT FILING DATE: 1996-11-22
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: Patent In version 4.1
? SEQ ID NO 4
? LENGTH: 405

```


Best Local Similarity 25.9%, Pred No. 13, 05,
Matches 79; Conservative 45; Mismatches 140; Indels 51; Gaps 12;

QY 4 ITARIGEPVLKCKGCAKPKPKUKEWIKNTQKFAWKVLSPOGG-----PW 50
Db 31 VNSAVGDFIIPCK---LDVQNLMP-----GK---WKYKPDGSPVFIARSTKKSVOY 80
QY 51 LSV-----ARVLENSLIFAVVILQNGIFPQAMNPNKETSNY-----RVRYQIP 99
Db 81 DVMPEYKDKMLSENYSILSNARISDEKKEVCMV-----TEDNVFIAPITIVKFKOP 134
QY 100 GKPEIVDSASELTAGVPKVKGTGCVSEGSYPAGTILSMWLDGKPLVPNEKGVSVKEGTFRHP 159
Db 135 SKPEIVSKALFETEQLKGLDGLSEUSYFDGNIWYRNGKVLHPLLEGAVVIFKEMDP 194
QY 160 FGIPTLQSLQSLVTPARGDPRPTFSCS---FSGNLPBPRALPTADIQIRVWE-VVFLPE 215
Db 195 VTQLYMTSTLEYTKK-ADIQMPTGCVTYGYS-----GQKIHSEQAVFDIYYPTEQ 248
QY 216 VGLVVEPEBGAAPAGTIVTILCEVEALFSPQIHWKMGVPELFPSPVLLILPEIGPQDQ 274
Db 249 VTIOVLPPKNAIKKGDNIITLKCLONGNPPPEEFLVLPQGEIHRSSNIYITLMDVRNAT 308
QY 275 GTYSC 279
Db 309 GDYKC 313

RESULT 11
US-09-808-602-70
; Sequence 70, Application US/0908602
; Publication No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkels, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mexes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115a1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 1566-647 CJP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-808-602-70

Query Match 11.4%, Score 201; DB 9; Length 793;
Best Local Similarity 23.4%, Pred No. 0.00022;
Matches 97; Conservative 49; Mismatches 131; Indels 148; Gaps 19;
QY 6 ARIGEPVLKCKGAPKPKPORLEWKUN-----TGRTEAMKVLSPQGGPMD 51
Db 65 ANPGQPIVLGCK-VGQTPPVQVSWKNGAELPEGTHHTLLANGSLLIHHFRLQGGSPD 123
QY 52 -----SVAR----- 55
Db 124 EGDYECVAQNRKFLVSRKAKLQAATMSDFVHPVQAVIGERGVARFOCOIHLGPKPLIT 183
QY 56 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNCKETKSNVRV----- 95
Db 184 WEKNRVPDITDDERTYLLPFGVLTITGLRAEDSGIFHCVCASNIA:VRVSHGARLTVSGG 243
QY 96 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNCKETKSNVRV----- 95
Db 184 WEKNRVPDITDDERTYLLPFGVLTITGLRAEDSGIFHCVCASNIA:VRVSHGARLTVSGG 243

QY 96 ---YUIK-KPEIVDSASELTAGV-PNKVGTGCVSEGSYPAGTILSW-HLDGKPLVPNEKGV 150
Db 244 SGTYK---EPTILVGPENLTLTVHTAVLECVATGN-PRPIVSWSLDGRPI--GVGEGIO 297
QY 151 VKEOTRRHPGTGFTIQRSLMVLPAWGQIHPPTFSCSPSGIPFP-IPALPTADIPQPVWER 210
Db 298 VL-----GTGNLIISD---VTVOHSG---VVVC---AANRRPGTRVAKTA--QGRLLVQ 339
QY 211 VFLEEVOLVVEPEGGAAPGTVTVTCEVPAQPSQIHWKMGVPLPL-----PPSPV 263
Db 340 APAEFVQ---HPQISIRPAGTITAMFTCOAGGPPPHVTWIKNOVILCAGCHVRLKNNST 396
QY 264 LILPEIGPOQGTYSVATH---SSHG-----PQSRVAVSIIE 300
Db 397 LSTISGVPEDEAIYQCAENIAGSSQASARLTVLMAEGLPGPPRNVRAVSSSTE 451

RESULT 12
US-09-908-193-23
; Sequence 23, Application US/0908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKELS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-908-193-23

Query Match 11.4%, Score 201; DB 9; Length 793;
Best Local Similarity 23.4%, Pred No. 0.00022;
Matches 97; Conservative 39; Mismatches 131; Indels 148; Gaps 19;
QY 6 ARIGEPVLKCKGAPKPKPORLEWKUN-----TGRTEAMKVLSPQGGPMD 51
Db 65 ANPGQPIVLGCK-VGQTPPVQVSWKNGAELPEGTHHTLLANGSLLIHHFRLQGGSPD 123
QY 52 -----SVAR----- 55
Db 124 EGDYECVAQNRKFLVSRKAKLQAATMSDFVHPVQAVIGERGVARFOCOIHLGPKPLIT 183
QY 56 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNCKETKSNVRV----- 95
Db 184 WEKNRVPDITDDERTYLLPFGVLTITGLRAEDSGIFHCVCASNIA:VRVSHGARLTVSGG 243
QY 96 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNCKETKSNVRV----- 95
Db 244 SGTYK---EPTILVGPENLTLTVHTAVLECVATGN-PRPIVSWSLDGRPI--GVGEGIO 297

```

QY 151 VKQTRRHPEIGLFTLQSELMVITPARGGDRPTFSCSFSGLRHRALRTAPQPRWEP 210
Db 298 VL -----GTGNLLISD VTVQHSQ ---VVV ANPPTPVETA QSRVVQ 339
QY 211 VPLEEVQVVEPRANAVAPGTVTLTCEVPAQSPQIHMKKGVPLPL -----PPSPV 263
Db 340 APAEFVQ ---HPQISIRKPAFTIAMPTCAQGEPPHVTWLKNOGVLAGCHVRLKNNST 396
QY 264 LILPEIGPQAGTYSKVATH ---SSBS -----PQESRAVSISIE 300
Db 397 LSIISGVPEEALYQVAENIAISSQASRLVLWAEGLGPPRNVRAVSVSSTE 451

RESULT 14
US 09-800-198-59
: Sequence 59, Application US/09000108
: Publication No. US20040047816A1
: GENERAL INFORMATION:
: APPLICANT: Vetrinet, Cornie AM
: APPLICANT: Fernandes, Edma
: APPLICANT: Shimkets, Richard A
: APPLICANT: Berrmann, John L
: APPLICANT: Majumder, Kumud
: APPLICANT: Mishra, Vishva
: APPLICANT: Mezes, Peter S
: APPLICANT: Rastelli, Luca
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: 15966-697
: CURRENT APPLICATION NUMBER: US/09/800,198
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,596
: PRIOR FILING DATE: 2000-03-04
: NUMBER OF SEQ ID NOS: 98
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 59
: LENGTH: 794
: TYPE: PRT
: ORGANISM: Mus musculus
US 09-800-198-59

Query Match 11.4% Score 2017 DB 97 Length 794
Best Local Similarity 23.4% Pred. No. 0.00022
Matches 97 Conservative 49 Mismatches 131 Indels 148 Gaps 19

```

```

QY 6 ARIGPEPLVLKCKAPKKPPQPLEWFLN -----TORFENMKVLSFQSGGWED 51
Db 65 ANGGQPIVLGRK VETTFPQVQSWPKFNQALPPTHTHTILANGSTTHHPRFPGQSPSD 124
QY 52 -----SVAR ----- 55
Db 124 EGDYEVQNRPGLLYSRRKRLQATMSDFHVPQAVTSPFGGVAPFQVTHLPEKPLIT 184
QY 56 -----VLINGSLFLPAVHLEEGEIPQLAMNNKKEFASNTVEEV 95
Db 184 WEKNRPHDIDDERYTLIKGVLIQTGLRAPSGLPHFVASNIAVRVSHGAKRLVSGSG 243
QY 96 - YQIPGKPEIVDSASELTAGV PKKVGTCVSEGSYPASTLSW ILLAKKPLVPKRGVS 150
Db 244 SRTYK -- EPTILVLPDNTTIVGQAVILEVATUNFPFIVSWLELMGFFI GVEHTQ 297
QY 151 VKQTRRHPEIGLFTLQSELMVITPARGGDRPTFSCSFSGLRHRALRTAPQPRWEP 210
Db 298 VL -----GTGNLLISD VTVQHSQ ---VVV ANPPTPVETA QSRVVQ 339
QY 211 VPLEEVQVVEPRANAVAPGTVTLTCEVPAQSPQIHMKKGVPLPL -----PPSPV 263
Db 340 APAEFVQ ---HPQISIRKPAFTIAMPTCAQGEPPHVTWLKNOGVLAGCHVRLKNNST 396
QY 264 LILPEIGPQAGTYSKVATH ---SSBS -----PQESRAVSISIE 300
Db 397 LSIISGVPEEALYQVAENIAISSQASRLVLWAEGLGPPRNVRAVSVSSTE 451

```

```

RESULT 14
US 09-908-193-19
: Sequence 19, Application US/099908193
: Publication No. US20020192748A1
: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: APPLICANT: SHIMKETS, RICHARD A.
: APPLICANT: ZERHUSEN, RYAN
: APPLICANT: MALYANKAR, URIEL M.
: APPLICANT: PADIGARU, MORALITHARA
: TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY
: FILE REFERENCE: 21402-062
: CURRENT APPLICATION NUMBER: US/09/908,193
: CURRENT FILING DATE: 2001-07-18
: PRIOR APPLICATION NUMBER: 60/222,274
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: 60/221,650
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: 60/221,234
: PRIOR FILING DATE: 2000-07-27
: PRIOR APPLICATION NUMBER: 60/220,912
: PRIOR FILING DATE: 2000-07-26
: PRIOR APPLICATION NUMBER: 60/218,875
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/218,870
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/218,901
: PRIOR FILING DATE: 2000-07-18
: NUMBER OF SEQ ID NOS: 74
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 19
: LENGTH: 1252
: TYPE: PRT
: ORGANISM: Mus musculus
US 09-908-193-19

Query Match 10.9% Score 1925 DB 97 Length 1252
Best Local Similarity 26.4% Pred. No. 0.0011
Matches 73 Conservative 39 Mismatches 120 Indels 45 Gaps 11

```

```

QY 49 FWDGVARVLENSLFLFAVLTQDGLFEPQLAMN - PNKEFASNYEPV VTJDEEP 102
Db 183 PERSGLLPNCVLIQVIVQUSAGSTRCAVNSARFQSIALVALRSLGATKQD 242
QY 103 EIVDSASELTAVPKR VTTVPSRSPAGTISW ILLKGF LVENEKGVSVKQETRRHP 159
Db 243 VVIVAAENTVWSGQVWMEVAGADPTTFVSWVQDKKI 284
QY 160 KTGIFFL QSPMLVTPARGGDRPTFSCSFSGLRHRALRTAPQPRWEP 218
Db 285 EIDVIGRTPELLASG - PRGGVVTAKKERTRTGTAAGAAERVLAAALISA 339
QY 275 VVERKAVAFGTVTLTCEVPAQSPQIHMKKGVPLPLSHV - ILLPEIG 270
Db 340 PEALSRTHASTARVCHASGDFRHALHWHGTFLLKNGRVKVGAGASLVITQIG 495
QY 271 PDGQTSCTVATHSHSGPQESRAVSVSISIEEGEPT 307
Db 496 TADAGYQVAENGAGTAAGAAFLAVVVRKGLPSAPT 443

```

```

RESULT 15
US 09-908-193-20
: Sequence 20, Application US/09908193
: Publication No. US20020192748A1
: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: APPLICANT: SHIMKETS, RICHARD A.
: APPLICANT: ZERHUSEN, RYAN
: APPLICANT: MALYANKAR, URIEL M.
: APPLICANT: PADIGARU, MORALITHARA
: TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES ENCODED THEREBY
: FILE REFERENCE: 21402-062

```



```

; CURRENT APPLICATION NUMBER: US/09/908,193
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-908-193-20

Query Match      10.9%   Score 192.5,   DB 9,   Length 1253,
Best Local Similarity 26.4%   Pred. No. 0.0011:
Matches 73; Conservative 49; Mismatches 120; Indels 45; Gaps 11;

QY 49 PWDVAVVLPNGSLFLPAVLIQDRIPTQAMN----PNSKETKSNYPVY--VYQIPGKP 102
DB 183 PFSFPIIIINWVQIIIDVQSHAGSYKCVATNSAPQPSQFASLIVALRGSLFATPQQD 242
QY 103 EIVDSASELTAGVPNK--VGVTVSEGSYPAGTLSW-HLDGKPLVPNEKGVSVKEQTRRH 159
DB 243 WIVAAAPENTIVVSGGVVMEKVASAD-ETPEFVSVVRQDKPI-----284
QY 160 ETGIFTI-OSELMVTARGCDPRPTSCSFSPCLPHHRAIRTAIPQRYWEPVPLFEVOL 218
DB 285 STEVIVLPTNLIILIASQ----PHISGVYVCRANKPTRENTAAALPVLAAPAISQA-- 339
QY 219 WVEPEGAVAPGSHVTLIGEVPAQPSQIHMMDGVPLPLPSPV-----LILPEIG 270
DB 340 ---PEALSRTRASTARFVCKRACGPRPALHLDGIPL--RPNGRVKKVQGGGSLVITQIG 395
QY 271 PQDQGYSCVATHSHGQVQESRAVSIISIEPGEEGPT 307
DB 396 LODAGYYQCVAAENSAGTACAAAPLAVVVRGLPSAPT 432

```

Search completed: May 30, 2003, 15:59:08
Job time : 28.1712 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:49:16 : Search time 38.1351 seconds
(without alignments)
836,935 million cell updates/sec

Title: US-09-872-185b-2

Perfect score: 1766

Sequence: 1 AONTARIGEPLVKCKGAP.....GSGLGTLALAIILGLATA 332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1766	100.0	404	161506	advanced glycosyla
2	1463	82.8	416	A42879	advanced glycosyla
3	1381	78.2	402	709062	probable advanced
4	219.5	12.4	583	139428	alcam human
5	217.5	12.3	645	198240	cell surface glyco
6	217	12.3	1344	T14316	right protein mo
7	212.5	12.0	1523	450478	actinin - gsdls
8	209	11.8	1612	730805	cell adhesion prot
9	208	11.8	1033	519247	cell adhesion prot
10	207.5	11.7	588	A45254	surface glycoprote
11	205.5	11.6	588	JH0506	adhesion molecule
12	203	11.5	1651	714160	transmembrane rece
13	199.5	11.3	587	JH0464	PM GPAS precursor
14	198	11.2	4391	A38096	perlecan precursor
15	194.5	11.0	584	150419	s-glycerin precurs
16	190.5	10.8	702	A36310	cardiac embryonic a
17	189.5	10.7	423	729640	hypothetical prote
18	189.5	10.7	1273	T43406	Ca
19	185.5	10.5	1521	834338	cell adhesion prote
20	185	10.5	3707	530452	cell adhesion prote
21	182	10.3	5175	720992	heparan sulfate pr
22	180	10.3	6104	T43200	hemophilic prote
23	180.5	10.2	1265	A37667	neural cell adhesi
24	179.5	10.2	459	741009	cell adhesion mole
25	177	10.0	6442	729457	protein HSP-89 - c
26	175.5	9.9	344	756553	neurotrophin rat
27	172.5	9.8	739	150476	vascular cell adhe
28	169.5	9.6	458	1WMSR1	cell adhesion mole
29	169.5	9.6	518	744024	cell adhesion mole

30	169.5	9.6	521	2	JC1508	biliary glycoprote
31	168	9.5	1051	2	A39712	kinase-like protei
32	165	9.3	1040	2	A49356	transient axonal g
33	164.5	9.3	1091	2	A58532	glial cell membran
34	164	9.3	1447	2	A54100	tumor suppressor p
35	162.5	9.2	588	2	I37202	B-CAM protein - hu
36	162	9.2	761	1	ICHUNG	neural cell adhesi
37	162	9.2	1070	2	JC4593	protein-tyrosine k
38	161	9.1	739	2	JN0581	vascular cell adhe
39	160	9.1	1427	2	I51669	-umor suppressor -
40	159.5	9.0	725	1	IJMSNG	neural cell adhesi
41	159.5	9.0	1115	1	IJMSNL	neural cell adhesi
42	159	9.0	1040	2	A4695	axonal glycoprotei
43	158.5	9.0	1357	1	A41060	neural cell adhesi
44	158.5	9.0	2222	2	T13924	sdh protein - frui
45	158	8.9	1268	1	A39640	neural cell adhesi

ALIGNMENTS

Result 1

161596

Advanced glycosylation end-products receptor precursor - human
N:Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprol.
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C:Accession: 161596; B42879; S27968
R:Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Miya, K.; Takahashi, E.; Ando, A.; Inoko, Genomics 23, 408-419, 1994
A:Title: Three genes in the human MHC class III region near the junction with the cln interpart of mouse mammary tumor gene int-3.
A:Reference number: A55562, MUID:95137587, PMID:7835890
A:Accession: 161596
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-404 <RES>
A:Cross-references: GR:028769; NID:9561657; FTDN:AAA03574.1; PID:g190846
R:Neuber, M.; Schmidt, A.M.; Hreff, J.; Yan, S.B.; Wang, F.; Pan, Y.C.; Elliston, K. J. Biol. Chem. 267, 14998-15004, 1992
A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation
A:Reference number: A42879; MUID:92340547, PMID:1378843
A:Accession: B42879
A:Molecule type: mRNA
A:Residues: 1-299, 101-404 <RES>
A:Cross-references: EMBL:M1211, NID:g190845; PID:AAA03574.1; PID:g190846
A:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBIP:109438)
A:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly cellular function, thus contributing to tissue lesions in diabetes.
C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide sites in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C:Genetics:
A:Gene: GPR-AGER
A:Cross-references: GDB:306354; OMT:600214
A:Map position: 6p21.3
A:Inserts: 18/1, 53/3, 119/1, 140/3, 170/2, 231/1, 274/3, 322/2, 331/1, 373/2
A:Function:
A:Description: neuronal receptor for angiotensin, a RNA-binding protein involved in ne C:Suprafamily: advanced glycosylation end products receptor, immunoglobulin homology C:Keywords: Alzheimer's disease, glycoprotein, receptor, transmembrane protein F:122/Domain: signal sequence #status predicted <SIG>
F:23/404/Domain: advanced glycosylation end products receptor #status predicted <SIG>
F:31/101/Domain: immunoglobulin homology <IM1>
F:137/220/Domain: immunoglobulin homology <IM2>
F:252/363/Domain: immunoglobulin homology <IM3>
F:345/362/Domain: transmembrane #status predicted <TM1>
F:363/404/Domain: intracellular #status predicted <IM1>
F:25/81/Binding site: carbohydrate (ASO) (covalent) #status predicted
F:38/39/144-208,255/301/disulfide bonds: #status predicted

Query Match: 100.0%, Score 1766; DB 1: Length 404;

QY 109 SELLACVINKVGLVSSAGSYFAGTISWRLQCKPLVFNKQ VSVKQIRRHPIETILIQ 167
 DB 156 AVGEP-AMVE-VIPEKRP ---EPITVWVWVYVLE- --EGKITIR 195
 QY 168 SELLMTDARGCUPRTFSNFSKGLPRRIALRTAPLQKWWPVLPEVQIVVPEGCA 226
 DB 196 GUKIMSHIFKSD AUMYKCVAS-NMAGERSCAA --ELVLEKPSFLRRPINO 246
 QY 227 VAKSGVTITGVFAVGEFQIHWKKGIVPLP- LPTSPVLLIPETIGPQAGTYSVA 281
 DB 247 VLADAVNLFCEVQGLQNLHKKKIDGKLPACRYETIRSDHSLIMIDVSSSEDEGTICVA 406
 QY 282 TRSHSGCPQSRVAVSISI 298
 DB 307 ENSVGRFAFASGSI SVHV 423
 RESULT 7
 150478
 neurotin - goldfish (fragment)
 C:Species: Carassius auratus (goldfish)
 C:Date: 13-Sep-1996 #sequence_revision 13 Sep-1996 #next_change 13 Sep-1996
 C:Accession: 150478
 Kilaassing, U.; Giordano, S.; Stecher, B.; Lottspeich, F.; Staermer, C.A.
 Differentiation 56, 21-29, 1994
 A:Title: Molecular characterization of fish neurotin, a growth associated cell surface protein DM-GRASP/sc-17BEN.
 A:Reference number: 150478; MUID:94299040; PMID:8026643
 A:Accession: 150478
 A:Status: preliminary; translated from GB/EMBL/Genbank
 A:Molecule type: mRNA
 A:Residues: 1523 <LAE.
 A:Cross-references: GB:L25056; NID:q407418; PID:q407419

Query Match 12.8%; Score 212.5; DB 2; Length 523;
 Best Local Similarity 25.5%; Pred. No. 8 3e 07;
 Matches 76; Conservative 40; Mismatches 133; Indels 49; Gaps 11;
 QY 9 GEHLVLRKCAKRRKQPLEWYLNFGTEAWKVLSPQSGPMSVAR 55
 DB 2 GEHLVVRKGGKIKGLGFLK ---WKYVA LKSGPGLLVKQAQAEIVASLD 52
 QY 56 ---VLNGLSELPVAGIQHGLFRCOAMNRNKEFKISNYRVVYVLDCKPQIVISA 108
 DB 53 GYKSRVSIANSLLIARGLAQVFTVMVVSFTNLEYS-VEVKVHKFSAPVIRKNA 111
 QY 109 SELLACVINKVGLVSSAGSYFAGTISWRLQCKPLVFNKQ VSVKQIRRHPIETILIQ 167
 DB 112 KELECKLTQIGCVVENANPADIWKKNQGLLVLDCKIITLSILKDKLIGLSSTSS 171
 QY 169 ELMVTPARKGUPRTFSNFSKGLPRRIALRTAPLQKWWPVLPEVQIVVPEGCA 226
 DB 172 KQYV-ARKHNSVQPTCF AKH VNGPIQ VSPESPHHYPTKYSQVVS 220
 QY 223 EGVAVARGTIVLVEVPAQSPQIHWKKGIVPLP- LPTSPVLLIPETIGPQAGTYSVA 281
 DB 221 Q SPIDREEDVTLKQACENPPTSPNINIKKKVTVITKLVYITVATVAVASGLYK 277

RESULT 8
 140805
 dnt1 protein mouse
 N:Alternate names: transmembrane receptor protein Rehol homolog
 C:Species: Mus musculus (house mouse)
 C:Date: 22 Oct 1999 #sequence_revision 22 Oct 1999 #next_change 22 Oct 1999
 C:Accession: 140805
 Kow, M.C.; Iow, N.; Fordham, R.; Rabbits, P.
 Submitted to the EMBL Data Library, July 1998
 A:Description: The mouse homologue of human DNT1/H rehol gene; protein sequence and chr
 A:Reference number: 220879
 A:Accession: 140805
 A:Status: preliminary; translated from GB/EMBL/Genbank

A:Molecule type: mRNA
 A:Residues: 11612 <WIM.
 A:Cross-references: EMBL:Y17794; NID:q429714; PID:q429714; PIR:q429714; PIR:q429714
 A:Experimental source: Brain
 C:Genetics:
 A:Gene: dnt1
 A:Map position: 16
 Query Match 11.8%; Score 209; DB 2; Length 1612;
 Best Local Similarity 25.4%; Pred. No. 6 2e 06;
 Matches 78; Conservative 44; Mismatches 144; Indels 52; Gaps 12;
 QY 9 GEHLVLRKCAKRRKQPLEWYLNFGTEAWKVLSPQSGPMSVAR VSVKQIRRHPIETILIQ 167
 DB 43 GEPATLNCK AESRPTPTIEWKGGIEVETIKD DPR SUPMLISGSLFLFLIV 94
 QY 68 ---GTLQPLPRPCAMNPNCKFTKSNYPVAVYVLPKFDVIVSASLTAQVIN KYTP 122
 DB 95 HCKSRPRPRVYICVARNYAEAVSHNASLEVAL RDEKFNPSVAVAVGIPAVMGC 192
 QY 123 VSGSYFAGTISWRLQCKPLVFNKQ VSVKQIRRHPIETILIQ SELLMTDARGCUPRTFS 182
 DB 153 QPFGCHPEPTISWKKDSPLQKLEKIRIR --- GKKMLITYTKSD 195
 QY 183 IFSCSFSGILRRKALRTAPLQKWWPVLPEVQIVVPEGCA 226
 DB 196 --- AKYV-VGIVNMVGERSEVALLVLEKSFVKKQSNLAVTVLSAEKCEA 246
 QY 240 PAQSPQIHWKKGIVPLP LPTSPVLLIPETIGPQAGTYSVA THSHSGCPQSRVAVSISI 298
 DB 237 EGVAVARGTIVLVEVPAQSPQIHWKKGIVPLP- LPTSPVLLIPETIGPQAGTYSVA 281
 QY 295 SISIEP 401
 DB 405 TLIVQEP 411
 RESULT 9
 519247
 cell adhesion protein gp160 Dfrc fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 13-Jan-1995 #sequence_revision 13 Jan 1995 #next_change 24 Sep 1999
 C:Accession: 519247
 Kapulido, D.; Campuzano, S.; Koda, T.; Modolell, J.; Barbacid, M.
 EMBO J. 11, 491-404, 1992
 A:Title: Dfrc, a Drosophila gene related to the tyk family of neurotrophin receptors.
 A:Reference number: 519247; MUID:92164624; PMID:1471458
 A:Accession: 519247
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 11033 <DOL.
 A:Cross-references: EMBL:X63453; NID:q7883; PIR:q7883; PID:q7883
 C:Genetics:
 A:Gene: Flybase:DFRC4B
 A:Cross-references: Flybase:DFRC4B004849
 C:Species: Flybase:DFRC4B004849
 C:Keywords: ATP
 E:53-526, Locatin. Protein kinase homology -KIN-
 E:538-706/Region: protein kinase ATP binding motif

Query Match 11.8%; Score 208; DB 2; Length 1044;
 Best Local Similarity 22.4%; Pred. No. 3 6e 06;
 Matches 88; Conservative 58; Mismatches 199; Indels 108; Gaps 18;
 QY 12 EVLEK-EGAP KEFFPELPEWLNLT-PTLEAWKVLSPQSGPMSVAR VSVKQIRRHPIETILIQ 167
 DB 133 LALKCHVESASDEPLEFTHWNSKLSLWKNVQ --- LQHRLIFRQ 179
 QY 69 IOPGGLPRCAMNPNCK-EKSNYPVAVYVLPKFDVIVSASLTAQVIN KYTP 192
 DB 180 SEIDGYPRTASNAAGPVMKSGYV-VYSSVYVLPPTPPKPNVPMESWIKQIFLPE 246
 QY 193 --- ELVLEKSLTACVPRKVG TCVHRSYVACI LSWHLD 148

DB 237 KRCAAGLEPAIAAPNHLRIVQGFPGQGIKGEHTALCTIYELPDELKAKR(I)LRKKD 296
 QY 139 GKPLVPNEKGV-----VKQTR--RHPETGLTLOSELMTVPAPGGDPP 181
 DB 297 GKILRQVFLGSGAPPCHSDSKDALLBEDAKVLHKONG--TLSPASTIASDAG----- 350
 QY 182 PTSCSPSPGLPRHPALPTAPIQPVVRRVPLFVLPVVEPREGAVANGTIT LCEVP 240
 DB 351 -QVQCQLQ--LEAHAPINSPPGLEV-----EQLKFPQPTSKNLELAVVAKVHCRAQ 402
 QY 241 AQSPQIHMKKGVFLP-----PSPVILPEIPQVQVYSCVAVHSSHGTOESRAV 294
 DB 403 GILPVPVQVVKUGENTTIPHVVEDANCTLFRNVNSEHPGNYTCLATNSOQINATVAI 462
 QY 295 SIIIEGEGGTAGSVGGSLGLTALALGILG 327
 DB 463 NV-VVTPKFSVPVPGPIETSKQGTVMHQAIC 494
 RESULT 10
 A45254
 surface glycoprotein BEN precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 27 Jun-1994 #sequence_revision 27-Jun-1994 #text_change 21-Jul-2000
 C:Accession: A45254; S19202
 R:Pourquie, O.; Corbel, C.; Le Gaer, J.P.; Bossier, J.; Le Douarin, N.M.
 A:Title: BEN, a surface glycoprotein of the immunoglobulin superfamily, is expressed in
 A:Reference number: A45254; MUID:92302224, PMID:1608932
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-588 <POU>
 A:Cross-references: EMBL:X64301; NID:363087; PIDN:CAA45579.1; EID:963088
 C:Keywords: glycoprotein

Query Match 11 7%; Score 207.5; DB 2; Length 588;
 Best Local Similarity 20.9%; Pred. No. 26-06;
 Matches 84; Conservative 46; Mismatches 147; Indels 125; Gaps 11;
 QY 4 ITARIQEPLVLRCKGAPKPPQPLFWK--LNTGPTTEAMKVLSPQGG----- 48
 DB 37 VNAVYGDITMPC-----RLEVPGDLMFGK---WKYEMPNCSPVFIAPRSSTKKNV 84
 QY 49 PWDGV-----ARVLPGSLFLPAVGIDEGIFRCQAMNNGKETKSNRYRVVYQIPGKP 102
 DB 85 QYDDVPDYKDRLSLSENYTLISKNARISDEKRFVCMVTEDDVSEPTV-VKVPKQSP 143
 QY 103 EIVDSASELTAGVKNKVCISGSPAGTILSWHLDDKPLVPNEKGVSKQTRRHPETG 162
 DB 144 EILHQADFLETEKLMKLGCVVRSYPEGNTVYKNGRVLPVEEVVILNRKYENRSTG 203
 QY 163 LFTLOSELMTVPARGDPRPTSCSPGLPRHRLARTAPIQPVWE--PVPLEEVQIWE 221
 DB 204 LFTMTSSLOYMPTK-EDANAKFTCIIVTVHGPSGQ--KTIQSEPVVDVHYVTEKVTIRVL 260
 QY 222 PEGGAVAPGTVTITCEVPAQSPQ-----LHMWKGVPPLPSPVILPEICPQOGT 276
 DB 261 SQSSTIKEDUNVILKCSGNGNPPPOFFLEYIPGETEGIRSSDVTMTDVRNATGEYKCS 320
 QY 247 -----LHMWKGVPPLP 258
 DB 321 LIDKSNMDDTTITVHYLDLQTPSGEVTQIGELPVSCTISSSRNATVFWIKDTRMKT 380
 QY 259 PPSVILPEIPGQDQVYSCVATHSHSGPOESRAVSISITIF 300
 DB 381 SPS-----FSSLYQVQAGNYICFTTHKEVEGLKKRKLKLV 418

RESULT 11

JH0506
 adhesion molecule SCI precursor - chicken

C:Species: Gallus gallus (chicken)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
 C:Accession: JH0506; PS0270
 R:Tanaka, H.; Matsui, T.; Agata, A.; Tomura, M.; Kubota, I.; McFarland, K.C.; Kohr, B.
 A:Title: Molecular cloning and expression of a novel adhesion molecule, SCI.
 A:Reference number: JH0506; MUID:92030150; PMID:1931049
 A:Accession: JH0506
 A:Molecule type: mRNA
 A:Residues: 1-588 <TAN>
 A:Cross-references: GB:S63376; NID:9238000; PIDN:AAR20170.1; PID:9238001
 A:Experimental source: embryo
 A:Accession: PS0270
 A:Molecule type: protein
 A:Residues: 34-48 <TANI>
 C:Comment: This protein is un-queily and transiently expressed on spinal cord motoneur
 C:Keywords: glycoprotein; transmembrane protein
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:34-588/Product: adhesion molecule SCI #status predicted <ADH>
 F:500-523/Domain: transmembrane #status predicted <TRA>
 F:101-173,199,271,312,366,462,485,504/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 11.6%; Score 205.5; DB 2; Length 588;
 Best Local Similarity 24.8%; Pred. No. 2.8e-06;
 Matches 75; Conservative 46; Mismatches 135; Indels 47; Gaps 11;
 QY 4 ITARIQEPLVLRCKGAPKPPQPLFWK--LNTGPTTEAMKVLSPQGG----- 48
 DB 37 VNAVYGDITMPC-----RLEVPGDLMFGK---WKYEMPNCSPVFIAPRSSTKKNV 84
 QY 49 PWDGV-----ARVLPGSLFLPAVGIDEGIFRCQAMNNGKETKSNRYRVVYQIPGKP 102
 DB 85 QYDDVPDYKDRLSLSENYTLISKNARISDEKRFVCMVTEDDVSEPTV-VKVPKQSP 143
 QY 103 EIVDSASELTAGVKNKVCISGSPAGTILSWHLDDKPLVPNEKGVSKQTRRHPETG 162
 DB 144 EILHQADFLETEKLMKLGCVVRSYPEGNTVYKNGRVLPVEEVVILNRKYENRSTG 203
 QY 163 LFTLOSELMTVPARGDPRPTSCSPGLPRHRLARTAPIQPVWE--PVPLEEVQIWE 221
 DB 204 LFTMTSSLOYMPTK-EDANAKFTCIIVTVHGPSGQ--KTIQSEPVVDVHYVTEKVTIRVL 260
 QY 222 PEGGAVAPGTVTITCEVPAQSPQ-----LHMWKGVPPLPSPVILPEICPQOGT 276
 DB 261 SQSSTIKEDUNVILKCSGNGNPPPOFFLEYIPGETEGIRSSDVTMTDVRNATGE 316
 QY 277 YSC 279
 DB 317 YKC 319

RESULT 12

T14160

transmembrane receptor protein Robol - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14160
 R:Kidd, T.; Brose, K.; Mitchell, K.J.; Fetter, R.D.; Tessier-Lavigne, M.; Goodman, C.
 A:Title: Roundabout controls axon crossing of the CNS midline and defines a novel sub
 A:Reference number: Z17897; MUID:98117249; PMID:9458045
 A:Accession: T14160
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1651 <KID>
 A:Cross-references: EMBL:AF041082; NID:39111215; FID:g2811216; PIDN:AAC39960.1
 C:Function:
 A:Description: appears to function as the gatekeeper controlling midline crossing
 C:Keywords: transmembrane protein

Query Match

11.5%; Score 203; DB 2; Length 1651;
 Best Local Similarity 24.8%; Pred. No. 1.3e-05;
 Matches 76; Conservative 46; Mismatches 133; Indels 52; Gaps 12;


```

113 SISKVROJNAKIFDQVGAISQGV-----GSRTELATYKIPAPPEITPN-- 158
114 ELTA.VPN KVATVARENSYFATLSWHLFKEKLVFENEKVSXKEZTFPIETGL 163
115 --SAPALSNMIK AUCISENSFSPNIIWKNQDPIQFEDTKIITLVPF-SNGL 215
116 FHLGELAVITAPQDQKERTFSCFSFGLIKHALATAPIQRVWFV--FLEEVOLVVE 221
117 VIVATLFSKVP-EDENSLPHATVHWL--LQUMTEL-SPRVHVTVPYTEHVELKVA 271
118 FFGAVAGCVVTLICEVFAQPSNGHHMK-----DQVELFTSPVLLPELIGQD 273
119 INACIVKEDVUKVLQADGNPAVTSFFKRLGDSWQUMISLADINQVLMENHVSXSS 331
120 QTYSC 279
121 SGLYSC 337

```

```

F66 7.70 Binding site - befanon sulfate (Ser) (covalent) #status predicted
F86 4.41 7.57 4.21 40.2 4186 4294 4046 4168 Binding site: carbohydrate
F12995 3933 4179 Binding site: chondroitin sulfate (Ser) (covalent) #status
predicted

```



GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:37:16 ; Search time 19,4414 seconds

(without alignments)

708,288 Million cell updates/sec

Title: US-09-872-185B-2

Perfect score: 1766

Sequence: 1 AQTATATGPEIVLKKCAPGSHGATLALGILGLGTA 332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	404	1	Q15109 homo sapien
2	1463	82.8	416	1	PAGE_HUMAN
3	1391	78.8	402	1	PAGE_BOVIN
4	1383	78.3	403	1	PAGE_RAT
5	1245	12.7	583	1	PAGE_MOUSE
6	219.5	12.4	583	1	C166_MOUSE
7	217.5	12.3	646	1	MU18_HUMAN
8	213	12.1	564	1	C166_HUMAN
9	212.5	12.0	555	1	C166_BRARE
10	205.5	11.6	588	1	C166_CARAU
11	190.5	10.8	702	1	C166_CHICK
12	189	10.7	4393	1	PCBM_HUMAN
13	186.5	10.6	1914	1	KM15_HUMAN
14	185	10.5	515	1	PVRL_PIG
15	185	10.5	3707	1	PCBM_MOUSE
16	180.5	10.2	417	1	ICAS_FABIT
17	177	10.0	924	1	ICAS_HUMAN
18	175.5	9.9	444	1	NIP1_PAT
19	172.5	9.8	739	1	VICAL_RAT
20	172	9.7	1266	1	NCAC_CHICK
21	170	9.6	1447	1	PCP_MOUSE
22	169.5	9.6	521	1	CEAL_MOUSE
23	168.5	9.5	515	1	PVRL_MOUSE
24	168	9.5	1051	1	PKR7_HUMAN
25	168	9.5	1070	1	PKR7_HUMAN
26	165	9.3	1040	1	AXO1_HUMAN
27	164	9.3	1447	1	DEC_HUMAN
28	163	9.2	517	1	PVRL_HUMAN
29	163	9.2	761	1	HCAC_HUMAN
30	162	9.2	848	1	NCAL_HUMAN
31	162	9.2	1234	1	NPHN_RAT
32	161	9.1	739	1	VICAL_HUMAN
33	160	9.1	837	1	NCW2_MOUSE

RESULT 1

RAGE_HUMAN

ID RAGE_HUMAN STANDARD; PRT: 404 AA.

AC Q15109; Q15279; Q9Y3R3; Q9H2X7;

DI 01-NOV-1997 (Rel. 35, Created)

DI 01-NOV-1997 (Rel. 35, Last sequence update)

DI 15-JUN-2002 (rel. 41, Last annotation update)

DE Advanced glycosylation end product-specific precursor

DE (Receptor for advanced glycosylation end products).

DE AGER OR RAGE.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RC SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Lung;

RX MEDLINE=92340547; PubMed=1378843;

RA Nepper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,

RA Elliston K., Stein D., Shaw A.;

RT Cloning and expression of a cell surface receptor for advanced

RT glycosylation end products of proteins.*;

RL J. Biol. Chem. 267:14998-15004(1992).

RN [2]

RC SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=95137587; PubMed=7835890;

RA Sudaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,

RA Inoko H., Ikemura T.;

RT Three genes in the human MHC class III region near the junction with

RT products, PBX2 homeobox gene and a notch homolog, human counterpart

RT of mouse mammary tumor gene int-3.*;

RL Genomics 23:408-419(1994).

RN [3]

RC SEQUENCE FROM N.A. (ISOFORM 1).

RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,

RA Banta A., Spies T., Hood L.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RC SEQUENCE FROM N.A. (ISOFORM 1).

RA Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,

RA Yamamoto H.;

RT Molecular heterogeneity of the receptor for advanced glycation

RT endproducts.*;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

RC SEQUENCE FROM N.A. (ISOFORM 2).

RA Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,

RA Schuler A., Huber G.;

RT cDNA cloning of a novel secreted isoform of the human Receptor for

RT Advanced Glycation End products (RAGE) and characterization of cells

RT co-expressing cell-surface scavenger receptors and Swedish mutant

RT amyloid precursor protein.*;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [6]

RC SEQUENCE FROM N.A. (ISOFORM 1).

RE	TISSUE: Lung;
RA	Strasbourg B.
RN	Submitted (JAN-2022) to the EMBL/GenBank/DDBJ databases.
RP	[7]
KA	SEQUENCE OF 1 12 FROM N.A.
RT	Hudson B.L., Fellers T.S.
FT	"Novel polymorphisms in the receptor for advanced glycation end products (RAGE) gene."
RL	Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases
CC	- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED RATE IN DIABETES.
CC	- SUBCELLULAR LOCATION: Type 1 membrane protein (isoform 1).
CC	- Secreted (isoform 2).
CC	- ALTERNATIVE PRODUCTS: 2 ISIFORMS; 1 (SHOWN HERE) AND 2/RAGESB2;
CC	ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC	- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN LIKE V-TYPE DOMAIN.
CC	- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS.
CC	this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.sib.ch/aboutencyc/ or send an email to license@sib.sib.ch)
EMBL:	M41211; AAA03574.1;
DR	EMBL: D28769; BAA05958.1;
DR	EMBL: D89336; AAR47491.1;
DR	EMBL: AH046432; AA889469.1;
DR	EMBL: AJ133822; CA843108.1;
DR	EMBL: BC020669; AAI020669.1;
DR	EMBL: AF208289; AA335728.1;
DR	GeneW: ICNC; 420; AGER.
DR	MM: 600214;
DR	InterPro: IPR003006; Ig MH*
DR	InterPro: IPR003598; Ig C2*
DR	InterPro: IPR003600; Ig L1k*
DR	Plant: PF00047; Ig 2.*
DR	SMART: SM00410; Ig L1k*; 1.
DR	SMART: SM00408; IgC2; 1.
DR	PROSITE: PS00290; IG MH*. 1.
KW	Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
KW	Alternative splicing; Polymorphism.
FT	SIGNAL 1 22 POTENTIAL.
FT	CHAIN 23 404 ADVANCED GLYCOSYLATION END PRODUCT SPECIFIC RECEPTOR EXTRACTED FROM (POTENTIAL).
FT	DOMAIN 23 342 POTENTIAL.
FT	TRANSMEM 343 363 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 364 404 IG LIKE V-TYPE DOMAIN.
FT	DOMAIN 41 106 IG LIKE C2-TYPE DOMAIN 1.
FT	DOMAIN 147 215 IG LIKE C2-TYPE DOMAIN 2.
FT	DOMAIN 252 308 POTENTIAL.
FT	DISELF1D 38 99 POTENTIAL.
FT	DISELF1D 144 208 POTENTIAL.
FT	DISELF1D 259 401 POTENTIAL.
FT	CARBHYD 25 25 N LINKED (GLCNAC...) (POTENTIAL).
FT	CARBHYD 81 81 N LINKED (GLCNAC...) (POTENTIAL).
FT	DOMAIN 480 484 MISSING (IN ISOFORM 2).
FT	VARSPLIC 54 67 CYP11IPSPVLITPEIGIQGGTVSYVATHSSHGQPSRA VSSTIEPGEGPTAGSVGSGLTALALGIIGLGITPAAL LGVTLIWRRRGRGFERRKAPENPFEPFPIFNSEEPFAG ESSTGTP -> VSLERKAGTRRGANGRCILGLIRAGNSS RHPGLQPHGPSRVAHWGHIVAKAAPRRGEPPKKPGRC GAEIKTESVGCT (IN ISOFORM 2).
FT	VARSPLIC 275 404 U . R.
FT	VARIANT 100 100 ZEUU-VAR_011558.
FT	CONFLICT 1 1 M . G (IN REF. 1).

SQ SEQUENCE 404 AA; 42802 MW; 60564C4460 CONCEPT CRG94;

Query Match 100.0%, Score 1766; DB 1; Length 404;
Best local Similarity 100.0%; Pred. No. 8, R 107;
Matches 342; Conservative 1; Mismatches 0; Indels 0; Gaps 0

QY 1 AGNITARIGEPLVLRKCAKKPPORLEWKINTGRTEAMKVLSPOGGHDSVAVKVLNG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 23 AONTIATGEPLVLRKCAKKPVRLEWKNLNGRTPANKVLSPOAGGWISAVKVLNG 82
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 SLFLPAVGIGDRIETPRCGAMNNPKFTKSVPYRVYGIPIKRIEIVASSELTAQVNVKG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 SLFLPAVGIGDRIETPRCGAMNNPKFTKSVPYRVYGIPIKRIEIVASSELTAQVNVKG 142
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 TCVSGSYPACTLSWHLLKGLVFNKGVSKVQTRHHPTGLETLTOSIELMVTIPARGDDP 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 TCVSGSYPACTLSWHLLKGLVFNKGVSKVQTRHHPTGLETLTOSIELMVTIPARGDDP 202
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 183 RTTSSTSTSTEGFPHRALEGTAFLEFVWEVEPLEEVQLVVELEAWAVALKALVLTSEVP 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 REFTSCSPCLPFRHALRIAPIQPVREVPLEEVQLVVELEGAVALKALVLTSEVP 262
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 AQPSQPHMMKGVWPLDPSPVLIETPGHQAGTYSVAHSINGDQSFAVSISTIE 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 263 AQPSQPHMMKGVWPLDPSPVLIETPGHQAGTYSVAHSINPEESRAVESISTIE 322
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 PDEEPITAGSWGSGICHLALALGILGLGTA 342
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 324 PDEEPITAGSWGSGICHLALALGILGLGTA 364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2

RACE_BOVIN

ID RACE_BOVIN STANDARD; PRO: 416 AA.

AC Q28173;

DT 01 NOV 1997 (Rel. 35, created)

DT 01 NOV 1997 (Rel. 45, last sequence update)

DT 15 JUN 2002 (Rel. 41, last annotation update)

DE Advanced glycosylation end product specific receptor precursor
DE (receptor for advanced glycosylation end products)
DE ACER FOR RAGE.
DC Bos taurus (Bovine).
CC Fukuyama; Metzger; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovina; Bovidae;
OX Hovidae; Bovinae; Bos.
NPRI-taxid-99413;
RN 111

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

TISSUE Lung;

MEDLINE 92340547; PubMed 1494043;

Nucleot M. Schmidt A.M.; Brett J.; Yan S.D.; Wang Y.; Fan Y.C.; Ellington K.; Stern D.; Shaw A.J.

"Cloning and expression of a cell surface receptor for advanced glycosylation end products of proteins."

J. Biol. Chem. 267:14998-15004(1992).

FUNCTION: MEDIANE INTERACTIONS OF ADVANCED GLYCOSYLATION END PRODUCTS (AGE) THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED RATE IN DIABETES.

-1 SUBCELLULAR LOCATION: Type I membrane protein.

-1 TISSUE SPECIFICITY: ENDOTHELIAL CELLS.

-1 SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

-1 SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN LIKE V TYPE DOMAIN

-1 SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement ([see http://www.ebi.ac.uk/submit/announcements](http://www.ebi.ac.uk/submit/announcements)) or send an email to license@ebi.ac.uk.

```

DR EMBL: M91212; AAA03575.1; .
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00408; IGG2; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal
FT SIGNAL 1 22
FT CHAIN 23 416
FT DOMAIN 23 352
FT TRANSMEM 353 373
FT DOMAIN 374 416
FT DOMAIN 31 105
FT DOMAIN 136 214
FT DOMAIN 262 318
FT DISULFID 38 98
FT DISULFID 143 207
FT DISULFID 269 311
FT CARBOHYD 25 25
FT CARBOHYD 80 80
FT DOMAIN 391 396
FT SEQUENCE 416 AA; 44182 MW; 8703815573767AE CRC64;
Query Match 82.8%; Score 1463; DB 1; Length 416;
Best Local Similarity 81.5%; Pred. No. 2.5e-87;
Matches 278; Conservative 21; Mismatches 30; Indels 12; Gaps 2;
QY 2 QNITARIGELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQGGPMDSVARVLPNGS 61
DB 24 QNITARIGKPLVNLCKGAPKPPQQLKNTGRTEAWKVLSPQ-GDPWDSVARVLPNGS 82
QY 62 LFLPAVGIDGESIFRCQAMNENKETSINRVVYQIPGKPEIVDSASELTAGVPNKVGT 121
DB 83 LLLPAVGIDGESIFRCQAMNENKETSINRVVYQIPGKPEIVDPASELMACVPNKVGT 142
QY 122 CVSEGSYFAGTSLWHLDCGLVNEKGVSVKEVTRHPETGLFTLQSELMTVPARGDPR 181
DB 143 CVSEGSYPAGTSLWHLDCGLVNEKGVSVKEETKPKTKGLTGLHSELMTVPARGALH 202
QY 182 PTFSCSPGLPRHRLARTAPIQPRVW-----EPVLEERVQIVVEPESGAVAPG 230
DB 203 PTFSCSPGLPRHRLARTAPIQPRVWSEHPGEGENVDVAVPLKVLVVEPESGAVAPG 262
QY 231 GTVTLTCEVAQSPQIHHMKGVVPLPSPVLLPEIGSPQATYSVATHSSHPQAE 290
DB 263 GTVTLTCEVAQSPQIHHMKGVVPLPSPVLLPEIGSPQATYSVATHSSHPQAE 322
QY 291 SRAVSTSTIEGEGPTASVWSGSGTSLALALGILRLAT 331
DB 323 SRAVSTSTIEGEGPTASVWSGSGTSLALALGILRLAT 363
RESULT 3
PAGE_PAT STANDARD; PRT; 402 AA.
AC Q63495;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Advanced glycosylation end product-specific receptor precursor
DE (Receptor for advanced glycosylation end products).
GN AGER OR RAGE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Lung;
RX MEDLINE-9716824; PUBMED-9221812.

```

Renard C, Chappey G, Wautier M.P., Nagashima M., Lundh E., Morser J., Zhao L., Schmidt A.M., Scherrmann J.M., Wautier J.L.; "Recombinant advanced glycation end product receptor pharmacokinetics in normal and diabetic rats."; Mol. Pharmacol. 52:54-62(1997).

CC FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED RATE IN DIABETES.

CC SUBCELLULAR LOCATION: Type I membrane protein.

CC TISSUE SPECIFICITY: ENDOTHELIAL CELLS.

CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: L33413; AAA2027.1; .

DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003598; Ig_C2.

DR Pfam: PF00047; Ig_3.

DR SMART: SM00410; Ig_Like; 1.

DR SMART: SM00408; IGG2; 1.

DR PROSITE: PS00290; IG_MHC; 1.

KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.

FT SIGNAL 1 22

FT CHAIN 23 402

FT DOMAIN 23 341

FT TRANSMEM 342 362

FT DOMAIN 363 402

FT DOMAIN 31 105

FT DOMAIN 136 213

FT DOMAIN 250 306

FT DISULFID 38 98

FT DISULFID 143 206

FT DISULFID 257 299

FT CARBOHYD 25 25

FT CARBOHYD 80 80

FT SEQUENCE 402 AA; 42663 MW; 594481BC3A51E94E CRC64;

Query Match 78.8%; Score 1391; DB 1; Length 402;

Best Local Similarity 79.5%; Pred. No. 1e-82;

Matches 263; Conservative 24; Mismatches 42; Indels 2; Gaps 2.

QY 2 QNITARIGELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQGGPMDSVARVLPNGS 61

DB 24 QNITARIGELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQ-3DPWDSVARVLPNGS 82

QY 62 LFLPAVGIDGESIFRCQAMNENKETSINRVVYQIPGKPEIVDSASELTAGVPNKVGT 121

DB 83 LLLPAVGIDGESIFRCQAMNENKETSINRVVYQIPGKPEIVDPASELMACVPNKVGT 142

QY 122 CVSEGSYFAGTSLWHLDCGLVNEKGVSVKEVTRHPETGLFTLQSELMTVPARGDPR 181

DB 143 CVSEGSYPAGTSLWHLDCGLVNEKGVSVKEETKPKTKGLTGLHSELMTVPARGALH 201

QY 182 PTFSCSPGLPRHRLARTAPIQPRVWEPVLEERVQIVVEPESGAVAPGVTTCVPA 241

DB 202 PTFSCSPGLPRHRLARTAPIQPRVWEPVLEERVQIVVEPESGAVAPGVTTCVPA 261

QY 242 QPSPIVHMKGVVPLPSPVLLPEIGSPQATYSVATHSSHPQESRAVSTSTIEP 301

DB 262 QPPPTHTWIKDGTPLPLAPSPVLLPEVGHEDGIYSCVATHPSHPQESPPVNRVET 321

QY 302 QEEGIAGSVGGSGLCGLALALGILGLGCLG 332

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; U95030; AAC06342.1; -.
DR EMBL; L25474; AAA37528.1; -.
DR HSSP; Q13740; 1KUC.
DR MGP; MG1; I313266; Alcam.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF06047; Ig; 3.
DR SMART; SM00409; Ig; 3.
DR SMART; SM00410; Ig_Like; 2.
DR PROSITE; PS00290; IG_MHC; FALSE_NKG.
KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 583
FT DOMAIN 28 527
FT TRANSIMP 528 549
FT DOMAIN 550 583
FT DOMAIN 36 120
FT DOMAIN 263 321
FT DOMAIN 347 399
FT DOMAIN 428 492
FT DISULFID 43 113
FT DISULFID 157 220
FT DISULFID 270 313
FT DISULFID 354 392
FT DISULFID 435 485
FT CARBOHYD 95 95
FT CARBOHYD 167 167
FT CARBOHYD 265 265
FT CARBOHYD 306 306
FT CARBOHYD 361 361
FT CARBOHYD 457 457
FT CARBOHYD 480 480
FT CARBOHYD 499 499
FT CONFLICT 427 454
FT CONFLICT 454 454
SQ SEQUENCE 583 AA: 65161 MW: 65161 MW: E7BAPAF8FCA8F9489 CRC64;

Query Match
Best local similarity 25.9% pred. No. 1.4e-07;
Matches 7q; conservative 46; Mismatches 129; Indels 51; Gaps 12;

QY 4 ITARIGEPVLVKGAPKPPQKLEWKLNTGRTENKWLSPQSG-----PW 50
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
31 VNSAYGDTIWPCKR---LDVQNLMP---GK---WKYEKPDGSPVFIAPRSSTKKSVOY 80
QY 51 DSVAP-----VLPNGSLFPAVGIDGIFRCQAMNRNGKETKSNV-----RVRYVQIP 93
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
81 DDVPEYKDRLSLSENVTLSTLANAKISDEKREVCMLV-----TEDNVFEAPTLVKYFKOP 134
QY 100 GKPEIVDSASFLTAGVPNKVCTCVSGSYDPAGTSLNHLGKPLVPNEKGVSKVKEQTRRP 159
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
135 SKPEIVKAPFLETDLKLGDCISRDSYDPGNITWRNGKVLQPVGEVAILFKKEIDP 194
QY 160 ETGLFTLQSELMTVPARGSDPRPTFSCS---FSPGLPRHRLRTAPIQPRVME-PVPLEE 215
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
195 CTQLYVTWTSSELYKFKTR-SDIQPFPTCSVTIYGPS-----GQKTIYSEQIFDIYVPTQ 248
QY 216 VQWVEPEGGAAGGVITLICEVPAQSP-QIHWKKGVPFLPSPVLLILPEIGPQDQ 274
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
249 VTIVLPPKNAIKEDNITLQCLNGNPPPEEFMFYLPQGEGRSSNTYTLTDVRNAT 308
QY 275 GTYSC 279
Db : : : : :
309 GDYKC 313

-----
EMBL; L38608; AAF59499.1; -.
EMBL; Y10183; CAA71256.1; -.
PDB; 1KJC; 03-APR-96.
Genew; HGNC:400; ALCAM.
MIM; 601662; -.
InterPro; IPR003599; Ig.
InterPro; IPR003606; Ig_MHC.
InterPro; IPR003600; Ig_Like.

```

```

DR 100047; 142 5;
DR SMART: SM00409; IG: 2;
DR SMART: SM00410; IG: 110; 2;
DR PRESITE: PS00290; IG: MUC; FALSE NEG.
KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal; 3D-structure; polymorphism.
FT SIGNAL 1 27
FT CHAIN 28 583
FT DOMAIN 28 583
FT TRANSMEM 528 549
FT DOMAIN 550 583
FT DOMAIN 36 120
FT DOMAIN 150 227
FT DOMAIN 263 321
FT DOMAIN 347 399
FT DOMAIN 428 492
FT DISULFID 43 114
FT DISULFID 157 220
FT DISULFID 270 313
FT DISULFID 454 392
FT DISULFID 435 485
FT CARBOHYD 91 91
FT CARBOHYD 167 167
FT CARBOHYD 265 265
FT CARBOHYD 306 306
FT CARBOHYD 461 461
FT CARBOHYD 457 457
FT CARBOHYD 480 480
FT CARBOHYD 499 499
FT VARIANT 258 258
FT VARIANT 401 401
FT VARIANT 401 401
FT SEQUENCE 583 AA: 65142 MW: 602487460284 CRC64:
Query Match 12.48; Score 219.5; DB: 1; Length 583;
Best Local Similarity 25.98; Pred. No: 30-07;
Matches 79; Conservative 45; Mismatches 130; Indels 51; Gaps 12;
QY 4 ITARIGELVLRKCAAPKPPGPIFWKINTGPTFAWVLSQGGV... ---PW 50
DB 31 VNSAVGQDITDTR...LVQGNMF...CK...WKYERKDGSIWVAFSSIKKSVQY 80
QY 51 DSV...ARVLNGLSLFVAVGQDSIFRCQAMNNGRETSNY...PVEVYQIP 99
DB 81 DAVPEYKORLINSNTYLLISNARSDEKREVMV...TELVPEAPTIVKVEKQP 134
QY 100 GKPEIVDSASLEFACVPAKKNVIVSSEGVPACTLSWHLGGAPVGNKAGSVKQIRKHP 159
DB 145 SKPEIVSNALFETFLPKIQLQISESYDGNITWYNGCVLHLEGVVITFKKMDP 194
QY 160 EIGLILSELNVIFARAGDERPFS...ESPCLPHRALPTAPLPVWE...PVPLEE 215
DB 195 VTQVMTISTEYKTKR...ADLMDPEVCSVTVYGS...QKTHSCVAFVIVYPTDQ 248
QY 216 VQLVVEPEGAVAPNSHTLTFEFAQSP...QIHWKIAVPLPLPSPVLLPGLPQDQ 274
DB 249 VTQVLPKRNALKEIDNITLKLNGNPPPEETFLYLGQPEGIRSSNTYTLMDVRRNAT 308
QY 275 GPYSC 279
DB 309 GUYK 314

```

RESULT 7

MUC18

ID MUC18_HUMAN

AC P43121

DT 01 NOV 1995 (rel. 32, created)

DT 01 NOV 1995 (rel. 42, last sequence update)

DT 15 JUN 2002 (rel. 41, last annotation update)

DE Cell surface glycoprotein MUC18 precursor (melanoma-associated antigen

```

DE MUC18 (Melanoma associated antigen A32) (35 endo, 1 endohectin)
DE Associated antigen (CD46 antigen) (Melanoma adhesion molecule).
DE MCAM OR MUC18.
DE Homo sapiens (Human).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DE NCBI_TaxID=9606;
DE [1]
DE SEQUENCE FROM N.A.
DE TISSUE-Melanoma;
DE MEDLINE-9009368; PubMed-2602481;
DE Lehmann J.M., Riethmuller G., Johnson J.P.:
DE "MUC18, a marker of tumor progression in human melanoma, shows
DE sequence similarity to the neural cell adhesion molecules of the
DE immunoglobulin superfamily."
DE Proc. Natl. Acad. Sci. U.S.A. 86:9891-9895(1989).
DE [2]
DE REVISIONS, SEQUENCE FROM N.A.
DE TISSUE-Melanoma;
DE MEDLINE-9331384; PubMed-8378324;
DE Sers C., Kirsch K., Rothbacher U., Riethmuller G., Johnson J.P.:
DE "Genomic organization of the melanoma associated glycoprotein MUC18:
DE implications for the evolution of the immunoglobulin domains."
DE Proc. Natl. Acad. Sci. U.S.A. 90:8514-8518(1993).
DE [3]
DE SEQUENCE OF 24-44; 98-112; 135-143; 240-260; 379-389 AND 460-478.
DE MEDLINE-94215196; PubMed-8162602;
DE Shih L.M., Elder D.E., Speicher D., Johnson J.P., Herlyn M.:
DE "Isolation and functional characterization of the A32 melanoma
DE associated antigen."
DE Cancer Res. 54:2514-2520(1994).
DE [4]
DE SEQUENCE OF 27-49; 98-112 AND 246-260.
DE MEDLINE-96136302; PubMed-857133;
DE Bardin N., Frances V., Lesaulle G., Hirschowski N., George F.,
DE Sempol J.:
DE "Identification of the S Endo 1 endothelial associated antigen."
DE Biochem. Biophys. Res. Commun. 218:210-216(1996).
DE [5]
DE FUNCTION.
DE MEDLINE 94132529; PubMed-829890;
DE Johnson J.P., Rothbacher U., Sers C.:
DE "The progression associated antigen MUC18: a unique member of the
DE immunoglobulin supergene family."
DE Melanoma Res. 3:337-340(1993).
DE -1- FUNCTION: COULD BE AN ADHESION MOLECULE ACTIVE IN NEURAL CREST
DE CELLS DURING EMBRYONIC DEVELOPMENT. ITS EXPRESSION MAY ALLOW
DE MELANOMA CELLS TO INTERACT WITH CELLULAR ELEMENTS OF THE VASCULAR
DE SYSTEM THEREBY ENHANCING HEMATOGENOUS TUMOR SPREAD.
DE -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
DE -1- TISSUE SPECIFICITY: MAY APPEAR AT THE SURFACE OF NEURAL CREST
DE CELLS DURING THEIR EMBRYONIC MIGRATION. APPEARS TO BE LIMITED TO
DE VASCULAR SMOOTH MUSCLE IN NORMAL ADULT TISSUES. ASSOCIATED WITH
DE TUMOR PROGRESSION AND THE DEVELOPMENT OF METASTASIS IN HUMAN
DE MALIGNANT MELANOMA. EXPRESSED MOST STRONGLY ON MELANOCYTIC LESIONS
DE AND ADVANCED PRIMARY TUMORS AND IS ONLY RARELY DETECTED IN BENIGN
DE MELANOCYTIC NEVI AND THIN PRIMARY MELANOMAS WITH A LOW PROBABILITY
DE OF METASTASIS.
DE -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
DE -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS.
DE -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE V TYPE DOMAINS.
DE -1- DATABASE: NAME=PRO; NOTE=CD guide CD46 entry
DE WWW: http://www.ncbi.nlm.nih.gov/efw/efw4cd46.htm
DE
DE This SWISS-PROT entry is copyright. It is produced through a collaboration
DE between the Swiss Institute of Bioinformatics and the EMBL outstation
DE at the European Bioinformatics Institute. There are no restrictions on its
DE use by non profit institutions as long as its content is in no way
DE modified and this statement is not removed. Usage by and for commercial
DE entities requires a license agreement! (See http://www.ebi.ac.uk/announcements/
DE or send an email to license@ebi.ac.uk)
DE
DE EMBL: M29277; AAA30824.1;

```


DR EMBL: M28882; AAA20922.1; ..
 DR EMBL: X68254; CAA48332.1; ..
 DR EMBL: X68265; CAA48332.1; JOINED.
 DR EMBL: X68266; CAA48332.1; JOINED.
 DR EMBL: X68267; CAA48332.1; JOINED.
 DR EMBL: X68268; CAA48332.1; JOINED.
 DR EMBL: X68270; CAA48332.1; JOINED.
 DR EMBL: X68271; CAA48332.1; JOINED.
 DR Genes: HGNC:6934; MCAM.
 DR MIN: 155735; ..
 DR InterPro: IPR003598; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_5.
 DR SMART: SM00410; Ig_Like; 2.
 DR SMART: SM00408; Ig_2; 2.
 KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
 KW Repeat; Signal;
 FT SIGNAL: 1 23
 FT CHAIN: 24 646 CHIL SURFACE GLYCOPROTEIN MUC1R
 FT DOMAIN: 24 559 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 560 583 POTENTIAL.
 FT DOMAIN: 584 646 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN: 41 123 IG-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN: 154 230 IG-LIKE V-TYPE DOMAIN 2.
 FT DOMAIN: 255 327 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN: 358 414 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN: 445 506 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN: 48 116 PROBABLE.
 FT DISULFID: 161 223 PROBABLE.
 FT DISULFID: 272 329 PROBABLE.
 FT DISULFID: 365 407 PROBABLE.
 FT DISULFID: 452 499 PROBABLE.
 FT CARBOHYD: 56 56 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 418 418 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 449 449 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 467 467 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 508 508 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 518 518 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 527 527 N-LINKED (GLCNAC: ..) (POTENTIAL).
 FT CARBOHYD: 544 544 N-LINKED (GLCNAC: ..) (POTENTIAL).
 SQ SEQUENCE: 646 AA, 71793 MW, F064A5DAER0BARC6 CRG64.
 Query Match: 12.3%; Score 217.5; DB 1; Length 646.
 Best Local Similarity: 23.3%; Pred. No. 4.5e-07;
 Matches: 95; Conservative: 55; Mismatches: 135; Indels: 123; Gaps: 19;
 QY 4 ITATIGPLVLKCKGAIKKPP-QLRW--KLNTGRTEAMKVLSPQG---GGPWDSVARVL 57
 DB 36 VEVHVGSTALKKGLSUSQGNLSHVIMF-SVHKKKTHIFPVYDQDQSSPFGYKRLSLD 95
 QY 58 PNR-SLELPVAVTIGRIFPVQMMNINRRTKS-NYPV--PVYQTPSRKEI-----V 105
 DB 96 DRGALLALQVLPQDRIKFIQO-----GRKPKSQRYRQLRYKAPENIQVNIPLIPV 150
 QY 106 DSASELTAGVNPVAVGTCVSGSYDPATITSWHLEGRPIVINEKGSVSKVQIPEPPHPEIET 165
 DB 151 NSKE-----PEEATVCGVSNYPPIQVVIWYNGRPLKEEKNRVI--USSQVSSGLYT 203
 QY 166 IQSHLVTPAGCDPRITSCSPICPLPHPALR-----TAPI---QPRVW ETVPL- 213
 DB 204 LQSLKAQLVK-EKUAQFYCELNYRUPSGNHMKESREVTVPVFPYPTKVLVEPVGML 262
 QY 214 -----EE----- 215
 DB 263 KEGDVPVIRCIANQNDPHPSISKQNPSTREAREETTINGVILVLEPAKHSRGYEQQA 322
 QY 216 -----VQLVVFPEP-----CAVAP-----GVVILRQVPAQSPQTHMKIKGV 254
 DB 323 WLLDTMLSLSEFQELLNVYVSVKVSFAAPFQFSSITLTCESSNQULEPFWLREET 362
 QY 255 PLPLPPSPVLLDEIGHQDGIYSQVAT-HSSHGPFQFSRAVSSITIEP 301

DR 383 LGVIFRSPVLEQLHHLKPPAAGSYKRVASVPSIPGIPNQLVVKLATFGP 440
 RESULT 8
 C166_RRAPE STANDARD; PRT: 564 AA.
 AC Q90460; Q90480;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C166 antigen homolog precursor (Neurotin) (DM-GRASP homolog).
 GN C166.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID:7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 PX MEDLINE:94376084; PubMed:80896660;
 RA Karki J.P., Chang S., Kuwada J.Y.;
 RT "The molecular cloning and characterization of potential chick
 DM-GRASP homologs in zebrafish and mouse.";
 RL J. Neurobiol. 25:831-845(1994).
 RN [2]
 RP SEQUENCE OF 398-561 FROM N.A.
 RX MEDLINE:94299340; PubMed:8026643;
 RA Laessing U., Giordano S., Strecher R., Iottspeich F., Stuermer C.A.;
 RT "Molecular characterization of fish neurotin: a growth-associated
 cell surface protein and member of the immunoglobulin superfamily in
 the fish retinotectal system with similarities to chick protein
 DM-GRASP/SC-1/HER.";
 RL Differentiation 56:21-29(1994).
 CC -|- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY
 NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -|- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -|- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).
 DR EMBL: L25273; AAA50024.1; ..
 DR EMBL: L25057; AAA50048.1; ..
 DR HSP: Q13740; IKIC;
 DR ZFIN: ZDB-GENE-990415-40; cal66
 DR InterPro: IPR003606; Ig_MHC.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_4.
 DR SMART: SM00410; Ig_Like; 3.
 DR SMART: SM00406; Ig_V; 1.
 DR PROSITE: PS00240; Ig_MHC; 1
 KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
 KW Repeat; Signal;
 FT SIGNAL: 1 24 POTENTIAL.
 FT CHAIN: 25 564 C166 ANTIGEN HOMOLOG.
 FT DOMAIN: 25 507 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 508 528 POTENTIAL.
 FT DOMAIN: 529 564 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN: 31 117 IG-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN: 147 222 IG-LIKE V-TYPE DOMAIN 2.
 FT DOMAIN: 256 314 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN: 340 389 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN: 419 477 IG-LIKE C2-TYPE DOMAIN 3.
 FT DISULFID: 38 110 POTENTIAL.
 FT DISULFID: 154 217 POTENTIAL.

FT DISULFID 263 306 POTENTIAL.
 FT DISULFID 426 470 POTENTIAL.
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 488 497 LFEEDKPKFG -> R (IN REF. 2).
 SQ SEQUENCE 564 AA: 61273 MW: ADAA7C293A607929 CRC64;
 Query Match 12.1% Score 213; DB 1; Length 564;
 Best Local Similarity 24.7% Pred No 7 5e-07;
 Matches 75; Conservative 52; Mismatches 144; Indels 46; Gaps 12;
 QY 9 GEPLVLCCKGAPKPPQLEWKINTQTEAWKVLSPQGGPMDVSAR ----- 55
 DB 31 GETIVPCNDCTKKPKDGLFTK - - - - -WK-YARDGSGPGLIKQAQKIDPIVVSAMD 81
 QY 56 -----VLNGSLFLPAVGLQDEGIFRCQAMNNGKETKSNYRVVQIIPCKPEIVDSA 108
 DB 82 GYKTRVSIANSLLIAQSLTQURVETCMVSSNTLE-EFSVEVKVHKKPSAPVKNKV 140
 QY 109 SELTAGVFNKVGICVSGSYPACTLSWHLJGKPLVFNKGVSKQKTRKHHPHPTGLTIQS 168
 DB 141 KELENGKLTQLGECVGSANPAADLIWMKNNQALVDDGKTLITSDVTKDPTVGLSSTSS 200
 QY 159 ELMTVPARGSLPPTFS-SFSP-SLPRIHALFTAPIQ PYWKE FVPLEEVLVVEPEG 224
 DB 201 KQVTF-ARKEDVASQFTIC-----VAKH-----VTGPNQVSTPDITFIHYPTKXSLQVVSQ- 250
 QY 225 GAVAFKGTIVTTEVPAQSP-QIHWKMGKVPPLPSPVLIIPFISGPDQGTYSQVATH 283
 DB 251 SPIREGDDVTLKQADGNPPPTSFENFNKGGKVTVDKQVYTIIGVTRADSGWYKC-SIL 309
 QY 284 SSHGQHSRAVSIISLIE 300
 DB 310 DNDVMESTQIVTWSFID 326
 RESULT 9
 ID C166 CARAU STANDARD; PRT: 555 AA.
 AC Q90304;
 DT 30-NOV-1997 (rel. 35, Created)
 DT 15-JUN-2002 (rel. 41, Last sequence update)
 DE C166 antigen homolog precursor (Neurolin) (DM-GRASP homolog).
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Laessle U.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 30-552 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Retina;
 RX MEDLINE=94299040; PubMed 8026543;
 RA Laessle U.; Giordano S.; Stecher B.; Lottspeich F.; Stuermer C.A.O.;
 PT "Molecular characterization of fish neurolin: a growth-associated
 cell surface protein and member of the immunoglobulin superfamily in
 the fish retinotectal system with similarities to chick protein
 DM GRASP/SC-1/HRN.";
 RI DM GRASP/SC-1/HRN.";
 RL Differentiation 36:21-29(1994).
 CC -1- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY
 CC -1- NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- TISSUE SPECIFICITY: PRESENT ON ALL RETINAL GANGLION CELLS (RGCs)
 CC AND THEIR AXONS (IN EMBRYO). ABSENT FROM MATURE AXONS ALONG MOST
 CC OF THEIR LENGTH, BUT IS PRESENT ON NEW AND GROWING AXONS DERIVED
 CC FROM THE RGCs AT THE RETINAL MARGIN. REMAINS ON ADULT RGCs ONLY AT

CC CELL-CELL CONTACT SITES AND IS CONTINUOUSLY EXPRESSED IN THE RETINAL
 CC AXON TERMINAL ARBOR LAYERS OF THE ADULT TECTUM.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2 TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyright . It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.sib.ch/announcements>
 CC or send an email to license@sib.ch).
 CC
 CC EMBL: L25056; AAC48015.2;
 DR HSHP: Q13740; IACJ;
 DR InterPro: IPR003006; Iq_MHC;
 DR InterPro: IPR003600; Iq_Like
 DR InterPro: IPR003596; Iq_V;
 DR Pfam: PF00047; Iq; 3;
 DR SMART: SM00410; Iq_Like; 3;
 DR SMART: SM00406; Iq; 1;
 DR PROSITE: PS00290; Iq_MHC; 1;
 DR Repeat: Glycoprotein; Transmembrane.
 KW Signal; Developmental protein; Cell adhesion; Immunoglobulin domain;
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 555 C166 ANTIGEN HOMOLOG;
 FT TRANSMEM 24 499 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 500 520 POTENTIAL.
 FT DOMAIN 521 555 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1 31 Iq-LIKE V-TYPE DOMAIN 1;
 FT DOMAIN 147 224 Iq-LIKE V-TYPE DOMAIN 2;
 FT DOMAIN 256 313 Iq-LIKE C2 TYPE DOMAIN 1;
 FT DOMAIN 340 392 Iq-LIKE C2 TYPE DOMAIN 2;
 FT DOMAIN 419 477 Iq-LIKE C2 TYPE DOMAIN 3;
 FT DISULFID 38 110 POTENTIAL.
 FT DISULFID 154 217 POTENTIAL.
 FT DISULFID 264 406 POTENTIAL.
 FT DISULFID 426 470 POTENTIAL.
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 555 AA: 60371 MW: 5448014F00HFF68 C6064;
 Query Match 12.0% Score 212.5; DB 1; Length 555;
 Best Local Similarity 25.5% Pred No 8e 07;
 Matches 76; Conservative 40; Mismatches 144; Indels 49; Gaps 11;
 QY 9 GEPLVLCCKGAPKPPQLEWKINTQTEAWKVLSPQGGPMDVSAR - - - - - 55
 DB 31 GETIVPCNDCTKKPKDGLFTK - - - - -WKYVK DQSTPDELVKQAKRQFAIVSAD 81
 QY 56 -----VLNGSLFLPAVGLQDEGIFRCQAMNNGKETKSNYRVVQIIPCKPEIVDSA 108
 DB 82 GYKTRVSIANSLLIAQSLTQURVETCMVSSNTLEYS VEVVHKKPSAPVKNNA 140
 QY 109 SELTAGVFNKVGICVSGSYPACTLSWHLJGKPLVFNKGVSKQKTRKHHPHPTGLTIQS 168
 DB 141 KELENGKLTQLGECVGSANPAADLIWMKNNQTLVDDGKTLITSDVTKDPTVGLSSTSS 200
 QY 159 ELMTVPARGSLPPTFS-SFSP-SLPRIHALFTAPIQFVWVF VILVLCVAVVF 222
 DB 201 KQVTF-ARKEDVASQFTIC-----AKH-----VMGPDQ VSEDSFTTHYTKVNSQVNS 249
 QY 225 EGGAVALPGITVLTCEVPAQSP-QIHWKMGKVPPLPSPVLIIPFISGPDQGTYSQ 279
 DB 250 Q-SPIREGDDVTLKQADGNPPPTSFENFNKGGKVTVDKQVYTIIGVTRADSGWYKC 306
 RESULT 10
 C166_CHICK

ID C166-CHICK STANDARD: PRI: 588 AA.
 AC P4226;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C166 antigen precursor (Swi glycoprotein) (BEN glycoprotein) (PM-
 DE GPASP protein) (JG7 protein).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eupholostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI:taxid:6041;
 UN [1]
 PN SEQUENCE FROM N A , AND SEQUENCE OF 34-53
 PP TISSUE=Embryo;
 RC MEDLINE=92010150; PubMed 1931849;
 RA Tanaka H., Matsui I., Agata A., Tomura M., Kubota I.,
 RA McFarland K.C., Kohr B., Lee A., Phillips H.S., Shelton D.L.;
 RA "Molecular cloning and expression of a novel adhesion molecule, Scl.,"
 RA Neuron 7:545-545(1991)
 RL [2]
 RP SEQUENCE FROM N A , AND SEQUENCE OF 34-53 AND 54-62
 RC MEDLINE=92402234; PubMed 1600032;
 RA Pourquie O., Corbel C., Le Car J.-P., Rossier J., Le Douarin N.M.;
 RA "HEN, a surface glycoprotein of the immunoglobulin superfamily, is
 RA expressed in a variety of developing systems,"
 RA Proc. Natl. Acad. Sci. U S A 89:5261-5265(1992)
 RN
 PP POSSIBLE FUNCTION.
 PP MEDLINE=92211411; PubMed 1313197;
 RA Pourquie O., Hallonet M.E.R., Le Douarin N.M.;
 RA "Association of BEN glycoprotein expression with climbing fiber
 RA axogenesis in the avian cerebellum,"
 RL J. Neurosci. 13:1548-1557(1992).
 CC -1- FUNCTION: HOMOPHILIC ADHESION MOLECULE ASSOCIATED WITH CLIMBING
 CC -1- FIBER AXONogenesis. SUPPORTS NEURITE EXTENSION.
 CC -1- SURCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A RESTRICTED POPULATION OF AXONS.
 CC WITHIN THE SPINAL CORD IT IS LOCALIZED TO AXONS IN THE DORSAL
 CC FUNICULUS, MIDLINE FLOOR PLATE CELLS, AND MOTONEURONS FOUND IN
 CC EPITHELIA, CENTRAL NERVOUS SYSTEM, AND HEMOPOIETIC CELLS. EACH
 CC DEVELOPING SYSTEM CARRIES DIFFERENTLY GLYCOSYLATED MOLECULAR FORMS
 CC OF BEN.
 CC -1- DEVELOPMENTAL STAGE: WIDELY EXPRESSED DURING EMBRYONIC
 CC DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.

 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMRL: S63276; AAB20170.1;
 CC EMBL: M76678; AAA48602.1;
 CC EMRL: X64301; CAA45579.1;
 CC HSP: Q13740; 1KJC.
 CC InterPro: IPR003599; Iq.
 CC InterPro: IPR003006; Iq_MHC
 CC InterPro: IPR003600; Iq_Like

DR PLAN: PR00047; Iq: 5;
 DR SMART: SM00409; Iq: 3;
 DR SMART: SM00410; Iq_Like: 2;
 DR PROSITE: PS00290; Iq_MHC; FALSE_NFG.
 KW Cell adhesion. Immunoglobulin domain, Glycoprotein, Transmembrane,
 KW Repeat: Signa...
 FT SIGNAL 1 33
 FT CHAIN 34 588
 FT DOMAIN 34 532
 FT TRANSMEM 533 553
 FT DOMAIN 554 588
 FT DOMAIN 42 126
 FT DOMAIN 136 233
 FT DOMAIN 259 326
 FT DOMAIN 352 404
 FT DOMAIN 433 497
 FT DISULFID 49 119
 FT DISULFID 163 226
 FT DISULFID 276 319
 FT DISULFID 359 397
 FT DISULFID 440 490
 FT CARBOHYD 101 101
 FT CARBOHYD 173 173
 FT CARBOHYD 196 199
 FT CARBOHYD 271 271
 FT CARBOHYD 312 312
 FT CARBOHYD 366 366
 FT CARBOHYD 462 462
 FT CARBOHYD 485 485
 FT CARBOHYD 504 504
 FT CONFLICT 1 10
 FT CONFLICT 25 25
 FT CONFLICT 112 113
 FT CONFLICT 329 329
 FT CONFLICT 401 402
 FT CONFLICT 409 409
 FT CONFLICT 421 421
 FT CONFLICT 439 439
 FT CONFLICT 457 457
 FT CONFLICT 475 475
 FT CONFLICT 493 493
 FT CONFLICT 511 511
 FT CONFLICT 529 529
 FT CONFLICT 547 547
 FT CONFLICT 565 565
 FT CONFLICT 583 583
 FT CONFLICT 601 601
 FT CONFLICT 619 619
 FT CONFLICT 637 637
 FT CONFLICT 655 655
 FT CONFLICT 673 673
 FT CONFLICT 691 691
 FT CONFLICT 709 709
 FT CONFLICT 727 727
 FT CONFLICT 745 745
 FT CONFLICT 763 763
 FT CONFLICT 781 781
 FT CONFLICT 799 799
 FT CONFLICT 817 817
 FT CONFLICT 835 835
 FT CONFLICT 853 853
 FT CONFLICT 871 871
 FT CONFLICT 889 889
 FT CONFLICT 907 907
 FT CONFLICT 925 925
 FT CONFLICT 943 943
 FT CONFLICT 961 961
 FT CONFLICT 979 979
 FT CONFLICT 997 997
 FT CONFLICT 1015 1015
 FT CONFLICT 1033 1033
 FT CONFLICT 1051 1051
 FT CONFLICT 1069 1069
 FT CONFLICT 1087 1087
 FT CONFLICT 1105 1105
 FT CONFLICT 1123 1123
 FT CONFLICT 1141 1141
 FT CONFLICT 1159 1159
 FT CONFLICT 1177 1177
 FT CONFLICT 1195 1195
 FT CONFLICT 1213 1213
 FT CONFLICT 1231 1231
 FT CONFLICT 1249 1249
 FT CONFLICT 1267 1267
 FT CONFLICT 1285 1285
 FT CONFLICT 1303 1303
 FT CONFLICT 1321 1321
 FT CONFLICT 1339 1339
 FT CONFLICT 1357 1357
 FT CONFLICT 1375 1375
 FT CONFLICT 1393 1393
 FT CONFLICT 1411 1411
 FT CONFLICT 1429 1429
 FT CONFLICT 1447 1447
 FT CONFLICT 1465 1465
 FT CONFLICT 1483 1483
 FT CONFLICT 1501 1501
 FT CONFLICT 1519 1519
 FT CONFLICT 1537 1537
 FT CONFLICT 1555 1555
 FT CONFLICT 1573 1573
 FT CONFLICT 1591 1591
 FT CONFLICT 1609 1609
 FT CONFLICT 1627 1627
 FT CONFLICT 1645 1645
 FT CONFLICT 1663 1663
 FT CONFLICT 1681 1681
 FT CONFLICT 1699 1699
 FT CONFLICT 1717 1717
 FT CONFLICT 1735 1735
 FT CONFLICT 1753 1753
 FT CONFLICT 1771 1771
 FT CONFLICT 1789 1789
 FT CONFLICT 1807 1807
 FT CONFLICT 1825 1825
 FT CONFLICT 1843 1843
 FT CONFLICT 1861 1861
 FT CONFLICT 1879 1879
 FT CONFLICT 1897 1897
 FT CONFLICT 1915 1915
 FT CONFLICT 1933 1933
 FT CONFLICT 1951 1951
 FT CONFLICT 1969 1969
 FT CONFLICT 1987 1987
 FT CONFLICT 2005 2005
 FT CONFLICT 2023 2023
 FT CONFLICT 2041 2041
 FT CONFLICT 2059 2059
 FT CONFLICT 2077 2077
 FT CONFLICT 2095 2095
 FT CONFLICT 2113 2113
 FT CONFLICT 2131 2131
 FT CONFLICT 2149 2149
 FT CONFLICT 2167 2167
 FT CONFLICT 2185 2185
 FT CONFLICT 2203 2203
 FT CONFLICT 2221 2221
 FT CONFLICT 2239 2239
 FT CONFLICT 2257 2257
 FT CONFLICT 2275 2275
 FT CONFLICT 2293 2293
 FT CONFLICT 2311 2311
 FT CONFLICT 2329 2329
 FT CONFLICT 2347 2347
 FT CONFLICT 2365 2365
 FT CONFLICT 2383 2383
 FT CONFLICT 2401 2401
 FT CONFLICT 2419 2419
 FT CONFLICT 2437 2437
 FT CONFLICT 2455 2455
 FT CONFLICT 2473 2473
 FT CONFLICT 2491 2491
 FT CONFLICT 2509 2509
 FT CONFLICT 2527 2527
 FT CONFLICT 2545 2545
 FT CONFLICT 2563 2563
 FT CONFLICT 2581 2581
 FT CONFLICT 2599 2599
 FT CONFLICT 2617 2617
 FT CONFLICT 2635 2635
 FT CONFLICT 2653 2653
 FT CONFLICT 2671 2671
 FT CONFLICT 2689 2689
 FT CONFLICT 2707 2707
 FT CONFLICT 2725 2725
 FT CONFLICT 2743 2743
 FT CONFLICT 2761 2761
 FT CONFLICT 2779 2779
 FT CONFLICT 2797 2797
 FT CONFLICT 2815 2815
 FT CONFLICT 2833 2833
 FT CONFLICT 2851 2851
 FT CONFLICT 2869 2869
 FT CONFLICT 2887 2887
 FT CONFLICT 2905 2905
 FT CONFLICT 2923 2923
 FT CONFLICT 2941 2941
 FT CONFLICT 2959 2959
 FT CONFLICT 2977 2977
 FT CONFLICT 2995 2995
 FT CONFLICT 3013 3013
 FT CONFLICT 3031 3031
 FT CONFLICT 3049 3049
 FT CONFLICT 3067 3067
 FT CONFLICT 3085 3085
 FT CONFLICT 3103 3103
 FT CONFLICT 3121 3121
 FT CONFLICT 3139 3139
 FT CONFLICT 3157 3157
 FT CONFLICT 3175 3175
 FT CONFLICT 3193 3193
 FT CONFLICT 3211 3211
 FT CONFLICT 3229 3229
 FT CONFLICT 3247 3247
 FT CONFLICT 3265 3265
 FT CONFLICT 3283 3283
 FT CONFLICT 3301 3301
 FT CONFLICT 3319 3319
 FT CONFLICT 3337 3337
 FT CONFLICT 3355 3355
 FT CONFLICT 3373 3373
 FT CONFLICT 3391 3391
 FT CONFLICT 3409 3409
 FT CONFLICT 3427 3427
 FT CONFLICT 3445 3445
 FT CONFLICT 3463 3463
 FT CONFLICT 3481 3481
 FT CONFLICT 3499 3499
 FT CONFLICT 3517 3517
 FT CONFLICT 3535 3535
 FT CONFLICT 3553 3553
 FT CONFLICT 3571 3571
 FT CONFLICT 3589 3589
 FT CONFLICT 3607 3607
 FT CONFLICT 3625 3625
 FT CONFLICT 3643 3643
 FT CONFLICT 3661 3661
 FT CONFLICT 3679 3679
 FT CONFLICT 3697 3697
 FT CONFLICT 3715 3715
 FT CONFLICT 3733 3733
 FT CONFLICT 3751 3751
 FT CONFLICT 3769 3769
 FT CONFLICT 3787 3787
 FT CONFLICT 3805 3805
 FT CONFLICT 3823 3823
 FT CONFLICT 3841 3841
 FT CONFLICT 3859 3859
 FT CONFLICT 3877 3877
 FT CONFLICT 3895 3895
 FT CONFLICT 3913 3913
 FT CONFLICT 3931 3931
 FT CONFLICT 3949 3949
 FT CONFLICT 3967 3967
 FT CONFLICT 3985 3985
 FT CONFLICT 4003 4003
 FT CONFLICT 4021 4021
 FT CONFLICT 4039 4039
 FT CONFLICT 4057 4057
 FT CONFLICT 4075 4075
 FT CONFLICT 4093 4093
 FT CONFLICT 4111 4111
 FT CONFLICT 4129 4129
 FT CONFLICT 4147 4147
 FT CONFLICT 4165 4165
 FT CONFLICT 4183 4183
 FT CONFLICT 4201 4201
 FT CONFLICT 4219 4219
 FT CONFLICT 4237 4237
 FT CONFLICT 4255 4255
 FT CONFLICT 4273 4273
 FT CONFLICT 4291 4291
 FT CONFLICT 4309 4309
 FT CONFLICT 4327 4327
 FT CONFLICT 4345 4345
 FT CONFLICT 4363 4363
 FT CONFLICT 4381 4381
 FT CONFLICT 4399 4399
 FT CONFLICT 4417 4417
 FT CONFLICT 4435 4435
 FT CONFLICT 4453 4453
 FT CONFLICT 4471 4471
 FT CONFLICT 4489 4489
 FT CONFLICT 4507 4507
 FT CONFLICT 4525 4525
 FT CONFLICT 4543 4543
 FT CONFLICT 4561 4561
 FT CONFLICT 4579 4579
 FT CONFLICT 4597 4597
 FT CONFLICT 4615 4615
 FT CONFLICT 4633 4633
 FT CONFLICT 4651 4651
 FT CONFLICT 4669 4669
 FT CONFLICT 4687 4687
 FT CONFLICT 4705 4705
 FT CONFLICT 4723 4723
 FT CONFLICT 4741 4741
 FT CONFLICT 4759 4759
 FT CONFLICT 4777 4777
 FT CONFLICT 4795 4795
 FT CONFLICT 4813 4813
 FT CONFLICT 4831 4831
 FT CONFLICT 4849 4849
 FT CONFLICT 4867 4867
 FT CONFLICT 4885 4885
 FT CONFLICT 4903 4903
 FT CONFLICT 4921 4921
 FT CONFLICT 4939 4939
 FT CONFLICT 4957 4957
 FT CONFLICT 4975 4975
 FT CONFLICT 4993 4993
 FT CONFLICT 5011 5011
 FT CONFLICT 5029 5029
 FT CONFLICT 5047 5047
 FT CONFLICT 5065 5065
 FT CONFLICT 5083 5083
 FT CONFLICT 5101 5101
 FT CONFLICT 5119 5119
 FT CONFLICT 5137 5137
 FT CONFLICT 5155 5155
 FT CONFLICT 5173 5173
 FT CONFLICT 5191 5191
 FT CONFLICT 5209 5209
 FT CONFLICT 5227 5227
 FT CONFLICT 5245 5245
 FT CONFLICT 5263 5263
 FT CONFLICT 5281 5281
 FT CONFLICT 5299 5299
 FT CONFLICT 5317 5317
 FT CONFLICT 5335 5335
 FT CONFLICT 5353 5353
 FT CONFLICT 5371 5371
 FT CONFLICT 5389 5389
 FT CONFLICT 5407 5407
 FT CONFLICT 5425 5425
 FT CONFLICT 5443 5443
 FT CONFLICT 5461 5461
 FT CONFLICT 5479 5479
 FT CONFLICT 5497 5497
 FT CONFLICT 5515 5515
 FT CONFLICT 5533 5533
 FT CONFLICT 5551 5551
 FT CONFLICT 5569 5569
 FT CONFLICT 5587 5587
 FT CONFLICT 5605 5605
 FT CONFLICT 5623 5623
 FT CONFLICT 5641 5641
 FT CONFLICT 5659 5659
 FT CONFLICT 5677 5677
 FT CONFLICT 5695 5695
 FT CONFLICT 5713 5713
 FT CONFLICT 5731 5731
 FT CONFLICT 5749 5749
 FT CONFLICT 5767 5767
 FT CONFLICT 5785 5785
 FT CONFLICT 5803 5803
 FT CONFLICT 5821 5821
 FT CONFLICT 5839 5839
 FT CONFLICT 5857 5857
 FT CONFLICT 5875 5875
 FT CONFLICT 5893 5893
 FT CONFLICT 5911 5911
 FT CONFLICT 5929 5929
 FT CONFLICT 5947 5947
 FT CONFLICT 5965 5965
 FT CONFLICT 5983 5983
 FT CONFLICT 6001 6001
 FT CONFLICT 6019 6019
 FT CONFLICT 6037 6037
 FT CONFLICT 6055 6055
 FT CONFLICT 6073 6073
 FT CONFLICT 6091 6091
 FT CONFLICT 6109 6109
 FT CONFLICT 6127 6127
 FT CONFLICT 6145 6145
 FT CONFLICT 6163 6163
 FT CONFLICT 6181 6181
 FT CONFLICT 6199 6199
 FT CONFLICT 6217 6217
 FT CONFLICT 6235 6235
 FT CONFLICT 6253 6253
 FT CONFLICT 6271 6271
 FT CONFLICT 6289 6289
 FT CONFLICT 6307 6307
 FT CONFLICT 6325 6325
 FT CONFLICT 6343 6343
 FT CONFLICT 6361 6361
 FT CONFLICT 6379 6379
 FT CONFLICT 6397 6397
 FT CONFLICT 6415 6415
 FT CONFLICT 6433 6433
 FT CONFLICT 6451 6451
 FT CONFLICT 6469 6469
 FT CONFLICT 6487 6487
 FT CONFLICT 6505 6505
 FT CONFLICT 6523 6523
 FT CONFLICT 6541 6541
 FT CONFLICT 6559 6559
 FT CONFLICT 6577 6577
 FT CONFLICT 6595 6595
 FT CONFLICT 6613 6613
 FT CONFLICT 6631 6631
 FT CONFLICT 6649 6649
 FT CONFLICT 6667 6667
 FT CONFLICT 6685 6685
 FT CONFLICT 6703 6703
 FT CONFLICT 6721 6721
 FT CONFLICT 6739 6739
 FT CONFLICT 6757 6757
 FT CONFLICT 6775 6775
 FT CONFLICT 6793 6793
 FT CONFLICT 6811 6811
 FT CONFLICT 6829 6829
 FT CONFLICT 6847 6847
 FT CONFLICT 6865 6865
 FT CONFLICT 6883 6883
 FT CONFLICT 6901 6901
 FT CONFLICT 6919 6919
 FT CONFLICT 6937 6937
 FT CONFLICT 6955 6955
 FT CONFLICT 6973 6973
 FT CONFLICT 6991 6991
 FT CONFLICT 7009 7009
 FT CONFLICT 7027 7027
 FT CONFLICT 7045 7045
 FT CONFLICT 7063 7063
 FT CONFLICT 7081 7081
 FT CONFLICT 7099 7099
 FT CONFLICT 7117 7117
 FT CONFLICT 7135 7135
 FT CONFLICT 7153 7153
 FT CONFLICT 7171 7171
 FT CONFLICT 7189 7189
 FT CONFLICT 7207 7207
 FT CONFLICT 7225 7225
 FT CONFLICT 7243 7243
 FT CONFLICT 7261 7261
 FT CONFLICT 7279 7279
 FT CONFLICT 7297 7297
 FT CONFLICT 7315 7315
 FT CONFLICT 7333 7333
 FT CONFLICT 7351 7351
 FT CONFLICT 7369 7369
 FT CONFLICT 7387 7387
 FT CONFLICT 7405 7405
 FT CONFLICT 7423 7423
 FT CONFLICT 7441 7441
 FT CONFLICT 7459 7459
 FT CONFLICT 7477 7477
 FT CONFLICT 7495 7495
 FT CONFLICT 7513 7513
 FT CONFLICT 7531 7531
 FT CONFLICT 7549 7549
 FT CONFLICT 7567 7567
 FT CONFLICT 7585 7585
 FT CONFLICT 7603 7603
 FT CONFLICT 7621 7621
 FT CONFLICT 7639 7639
 FT CONFLICT 7657 7657
 FT CONFLICT 7675 7675
 FT CONFLICT 7693 7693
 FT CONFLICT 7711 7711
 FT CONFLICT 7729 7729
 FT CONFLICT 7747 7747
 FT CONFLICT 7765 7765
 FT CONFLICT 7783 7783
 FT CONFLICT 7801 7801
 FT CONFLICT 7819 7819
 FT CONFLICT 7837 7837
 FT CONFLICT 7855 7855
 FT CONFLICT 7873 7873
 FT CONFLICT 7891 7891
 FT CONFLICT 7909 7909
 FT CONFLICT 7927 7927
 FT CONFLICT 7945 7945
 FT CONFLICT 7963 7963
 FT CONFLICT 7981 7981
 FT CONFLICT 7999 7999
 FT CONFLICT 8017 8017
 FT CONFLICT 8035 8035
 FT CONFLICT 8053 8053
 FT CONFLICT 8071 8071
 FT CONFLICT 8089 8089
 FT CONFLICT 8107 8107
 FT CONFLICT 8125 8125
 FT CONFLICT 8143 8143
 FT CONFLICT 8161 8161
 FT CONFLICT 8179 8179
 FT CONFLICT 8197 8197
 FT CONFLICT 8215 8215
 FT CONFLICT 8233 8233
 FT CONFLICT 8251 8251
 FT CONFLICT 8269 8269
 FT CONFLICT 8287 8287
 FT CONFLICT 8305 8305
 FT CONFLICT 8323 8323
 FT CONFLICT 8341 8341
 FT CONFLICT 8359 8359
 FT CONFLICT 8377 8377
 FT CONFLICT 8395 8395
 FT CONFLICT 8413 8413
 FT CONFLICT 8431 8431
 FT CONFLICT 8449 8449
 FT CONFLICT 8467 8467
 FT CONFLICT 8485 8485
 FT CONFLICT 8503 8503
 FT CONFLICT 8521 8521
 FT CONFLICT 8539 8539
 FT CONFLICT 8557 8557
 FT CONFLICT 8575 8575
 FT CONFLICT 8593 8593
 FT CONFLICT 8611 8611
 FT CONFLICT 8629 8629
 FT CONFLICT 8647 8647
 FT CONFLICT 8665 8665
 FT CONFLICT 8683 8683
 FT CONFLICT 8701 8701
 FT CONFLICT 8719 8719
 FT CONFLICT 8737 8737
 FT CONFLICT 8755 8755
 FT CONFLICT 8773 8773
 FT CONFLICT 8791 8791
 FT CONFLICT 8809 8809
 FT CONFLICT 8827 8827
 FT CONFLICT 8845 8845
 FT CONFLICT 8863 8863
 FT CONFLICT 8881 8881
 FT CONFLICT 8899 8899
 FT CONFLICT 8917 8917
 FT CONFLICT 8935 8935
 FT CONFLICT 8953 8953
 FT CONFLICT 8971 8971
 FT CONFLICT 8989 8989
 FT CONFLICT 9007 9007
 FT CONFLICT 9025 9025
 FT CONFLICT 9043 9043
 FT CONFLICT 9061 9061
 FT CONFLICT 9079 9079
 FT CONFLICT 9097 9097
 FT CONFLICT 9115 9115
 FT CONFLICT 9133 9133
 FT CONFLICT 9151 9151
 FT CONFLICT 9169 9169
 FT CONFLICT 9187 9187
 FT CONFLICT 9205 9205
 FT CONFLICT 9223 9223
 FT CONFLICT 9241 9241
 FT CONFLICT 9259 9259
 FT CONFLICT 9277 9277
 FT CONFLICT 9295 9295
 FT CONFLICT 9313 9313
 FT CONFLICT 9331 9331
 FT CONFLICT 9349 9349
 FT CONFLICT 9367 9367
 FT CONFLICT 9385 9385
 FT CONFLICT 9403 9403
 FT CONFLICT 9421 9421
 FT CONFLICT 9439 9439
 FT CONFLICT 9457 9457
 FT CONFLICT 9475 9475
 FT CONFLICT 9493 9493
 FT CONFLICT 9511 9511
 FT CONFLICT 9529 9529
 FT CONFLICT 9547 9547
 FT CONFLICT 9565 9565
 FT CONFLICT 9583 9583
 FT CONFLICT 9601 9601
 FT CONFLICT 9619 9619
 FT CONFLICT 9637 9637
 FT CONFLICT 9655 9655
 FT CONFLICT 9673 9673
 FT CONFLICT 9691 9691
 FT CONFLICT 9709 9709
 FT CONFLICT 9727 9727
 FT CONFLICT 9745 9745
 FT CONFLICT 9763 9763
 FT CONFLICT 9781 9781
 FT CONFLICT 9799 9799
 FT CONFLICT 9817 9817
 FT CONFLICT 9835 9835
 FT CONFLICT 9853 9853
 FT CONFLICT 9871 9871
 FT CONFLICT 9889 9889
 FT CONFLICT 9907 9907
 FT CONFLICT 9925 9925
 FT CONFLICT 9943 9943
 FT CONFLICT 9961 9961
 FT CONFLICT 9979 9979
 FT CONFLICT 9997 9997
 FT CONFLICT 10015 10015
 FT CONFLICT 10033 10033
 FT CONFLICT 10051 10051
 FT CONFLICT 10069 10069
 FT CONFLICT 10087 10087
 FT CONFLICT 10105 10105
 FT CONFLICT 10123 10123
 FT CONFLICT 10141 10141
 FT CONFLICT 10159 10159
 FT CONFLICT 10177 10177
 FT CONFLICT 10195 10195
 FT CONFLICT 10213 10213
 FT CONFLICT 10231 10231
 FT CONFLICT 10249 10249
 FT CONFLICT 10267 10267
 FT CONFLICT 10285 10285
 FT CONFLICT 10303 10303
 FT CONFLICT 10321 10321
 FT CONFLICT 10339 10339
 FT CONFLICT 10357 10357
 FT CONFLICT 10375 10375
 FT CONFLICT 10393 10393
 FT CONFLICT 10411 10411
 FT CONFLICT 10429 10429
 FT CONFLICT 10447 10447
 FT CONFLICT 10465 10465
 FT CONFLICT 10483 10483
 FT CONFLICT 10501 10501
 FT CONFLICT 10519 10519
 FT CONFLICT 10537 10537
 FT CONFLICT 10555 10555
 FT CONFLICT 10573 10573
 FT CONFLICT 10591 10591
 FT CONFLICT 10609 10609
 FT CONFLICT 10627 10627
 FT CONFLICT 10645 10645
 FT CONFLICT 10663 10663
 FT CONFLICT 10681 10681
 FT CONFLICT 10699 10699
 FT CONFLICT 10717 10717
 FT CONFLICT 10735 10735
 FT CONFLICT 10753 10753
 FT CONFLICT 10771 10771
 FT CONFLICT 10789 10789
 FT CONFLICT 10807 10807
 FT CONFLICT 10825 10825
 FT CONFLICT 10843 10843
 FT CONFLICT 10861 10861
 FT CONFLICT 10879 10879
 FT CONFLICT 10897 10897
 FT CONFLICT 10915 10915
 FT CONFLICT 10933 10933
 FT CONFLICT 10951 10951
 FT CONFLICT 10969 10969
 FT CONFLICT 10987 10987
 FT CONFLICT 11005 11005
 FT CONFLICT 11023 11023
 FT CONFLICT 11041 11041
 FT CONFLICT 11059 11059
 FT CONFLICT 11077 11077
 FT CONFLICT 11095 11095
 FT CONFLICT 11113 11113
 FT CONFLICT 11131 11131
 FT CONFLICT 11149

DE Carcinoembryonic antigen related cell adhesion molecule 5 precursor
DE (Carcinoembryonic antigen) (CEA) (Meconium antigen 100) (CD66e
DE ant body)
GN CEA/CA5 OR CEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE-90258861; PubMed-242441;
RA Schrewe H., Thompson J.E., Bona M., Hetta L.J.F., Maruya A.,
RA Bassauer M., Shively J.E., von Kleist S., Zimmermann W.,
RT "Cloning of the complete gene for carcinoembryonic antigen: analysis
RT of its promoter indicates a region conveying cell type-specific
RT expression.";
RF Mol. Cell. Biol. 10:2748-2748(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-88048876; PubMed-4670412;
RA Beaucourt N., Henschel S., Cournoyer D., Foks A., Stammers C.P.,
RI "Isolation and characterization of full-length functional cDNA clones
RI for human carcinoembryonic antigen.";
RL Mol. Cell. Biol. 7:3221-3221(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-89122014; PubMed-4220478;
RA Barnett T., Goebel S.O., Northrup M.A., Elting J.D.,
RI "Carcinoembryonic antigen family: characterization of cDNAs coding
RI for NCA and CEA and suggestion of nonrandom sequence variation in
RT their conserved loop domains.";
RL Genomics 3:59-66(1988).
RN [4]
RP SEQUENCE OF 5702 FROM N.A.
RX MEDLINE-87128144; PubMed-814146;
RA Oikawa S., Nakazato H., Kosaki G.,
RI "Primary structure of human carcinoembryonic antigen (CEA) deduced
RT from cDNA sequence.";
RL Biochem. Biophys. Res. Commun. 142:511-518(1987).
RN [5]
RP SEQUENCE OF 341 702 FROM N.A.
RX MEDLINE-87204247; PubMed-4034671;
RA Zimmermann W., Orthleib B., Friedrich R., von Kleist S.,
RI "Isolation and characterization of cDNA clones encoding the human
RI carcinoembryonic antigen reveal a highly conserved repeating
RI structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2960-2964(1987).
CC "SURFILLAP LOCATION: Attached to the membrane by a GPI anchor.
CC TISSUE SPECIFICITY: FOUND IN ADENOCARCINOMAS OF ENDOCRINALLY
CC DERIVED DIGESTIVE SYSTEM EPITHELIUM AND FETAL COLON.
CC PIM: COMPLEX IMMUNOREACTIVE GLYCOPROTEIN WITH A MW OF 180 kDa
CC COMPRISING 60% CARBOHYDRATE.
CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
CC SUBFAMILY.
CC SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN LIKE DOMAINS.
CC DATABASE: NAME-PROT: NOTE-CD guide (CD66e entry).
CC WWW="http://www.ncbi.nlm.nih.gov/seq/ncbi/cd66e.htm"

this SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/seq/ncbi/cd66e.htm>) or send an email to license@isb.ch).

EMBL: M17303; AA659513.1;
EMBL: M59262; AAA62845.1; ALT. SEQ.
EMBL: M59255; AAA62845.1; JOINED.
EMBL: M59256; AAA62845.1; JOINED.
EMBL: M59257; AAA62845.1; JOINED.
EMBL: M59258; AAA62845.1; JOINED.
EMBL: M59259; AAA62845.1; JOINED.

DR EMBL: M59260; AAA62845.1; JOINED.
DR EMBL: M59261; AAA62845.1; JOINED.
DR EMBL: M59709; NOT_ANNOTATED_CDS.
DR EMBL: M59710; NOT_ANNOTATED_CDS.
DR EMBL: M29540; AAA51967.1;
DR EMBL: X16455; CAA4474.1;
DR EMBL: M15042; AAA51963.1;
DR EMBL: M16234; AAA51972.1;
DR PIR: A36319; A36319.
DR Genew: HGNC:1817; CEA/CA5.
DR MIM: 114896;
DR InterPro: IPR004006; 1q_MHC.
DR InterPro: IPR004598; 1q_C2.
DR InterPro: IPR004600; 1q_Like.
DR Pfam: PF00047; 1q_6.
DR SMART: SM00410; IG_Like; 2.
DR SMART: SM00408; IGR2; 3.
KW Immunoglobulin domain; Glycoprotein; GPI anchor; Membrane; Signal;
KW Repeat.
FT SIGNAL 1 34
FT CHAIN 35 ?
FT PROPEP 2 702
FT DOMAIN 35 144
FT DOMAIN 146 237
FT DOMAIN 238 422
FT DOMAIN 324 415
FT DOMAIN 416 498
FT DOMAIN 502 543
FT DOMAIN 594 677
FT CARBOHYD 104 104
FT CARBOHYD 115 115
FT CARBOHYD 152 152
FT CARBOHYD 182 182
FT CARBOHYD 197 197
FT CARBOHYD 204 204
FT CARBOHYD 208 208
FT CARBOHYD 246 246
FT CARBOHYD 256 256
FT CARBOHYD 274 274
FT CARBOHYD 288 288
FT CARBOHYD 292 292
FT CARBOHYD 309 309
FT CARBOHYD 330 330
FT CARBOHYD 351 351
FT CARBOHYD 375 375
FT CARBOHYD 432 432
FT CARBOHYD 466 466
FT CARBOHYD 480 480
FT CARBOHYD 508 508
FT CARBOHYD 529 529
FT CARBOHYD 553 553
FT CARBOHYD 560 560
FT CARBOHYD 580 580
FT CARBOHYD 612 612
FT CARBOHYD 650 650
FT CARBOHYD 665 665
FT CONFLICT 320 320
FT SEQUENCE 702 AA, 6299AE26CD84C4C664;
Query Match 10.8%; Score 190.5; Db 1; Length 702;
Best local similarity 17.9%; Pred. No. 2,796; Pos.
Mat. Pos. 60; Conservative 62; Mismatches 144; Indels 211; Gaps 16;

QY 3 NITATIGEPVLKCKGAPKPPORLEWKLNTGTEAWKVLSPGGGHWISVARVLNGSL 62
Db 246 NISYRSKLENLSCHAA-SNPPAY-SNFVN
QY 63 FIPAVGIGQHPGEPGAMN-N-GKFTKSNTPVPVYVQIPKPIIVRSASE
Db 285 FIPNTVNNSSGYTCQAINSDTGLNKTVTITVVAEPKPKPFTSNNSPVEDEIAVLT 344

QY 111 -----LT-----AGVPNKVGTGCVS 124
 Db 345 CEPEIQNTYLVWVNGSLPVSRLQSLNDNRRTLLSVTRNDVGPYEGGIGNELSVDSHS 404
 QY 125 E- - - - -GSPYAGTLSMILLUCKPLVPNEKGV 150
 Db 405 DPVILNVLYGDDPTISPSVYYRPGVNLISLSCHAAASNPAAQYSLWIDG-----N 454
 QY 151 VKEQTR-----HPETGLFLLQSELMTVPARGD-----PRTFSCSPSP 190
 Db 455 IQQHTQELFTSNTIEKNSGLYTQAN---NSAGHSRTTIVTKITVSABLKPSSISNNNSK 511
 QY 191 GLPHRAL-----RTAPIQPRVW----- 208
 Db 512 PVEDKDAVAFTEPEAQTNTYLVWVNGSLPVSRLQSLNGNRRTLLFNVTNRDARAYVC 571
 QY 209 -----EVPLEEV-----QLVVEPEGNAVAPGGTVTLTCEVPAQSPQIHWMDG 253
 Db 572 GIONSVSANKSDPTLVLYGYDPTPIISPDDSSYLSGANILUSCHASNPSPQYSWRING 631
 QY 254 VPLPLPSPVLLLEFGPDQATYSVATHSHGHPQESRAVSISIIERPEGPTAGSVGG 313
 Db 632 I--POQHTQVLFIAKTIPNNNTYA--FVSNLATGRNNSIVKSTVSASGT-----SPGL 683
 QY 314 SGHCTLALALGILGGL 329
 Db 684 SAGNTVGIMIGLVGV 699

RESULT 12

PGBM_HUMAN

1D PGBM_HUMAN STANDARD; PRT: 4393 AA.
 AC p98160; Q16287;
 DT 01-OCT-1996 (Rel. 34, created)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Basement membrane-specific heparan sulfate proteoglycan core
 DE protein precursor (HSPG) (Perlecan) (PLC).
 GN HSPG2

OS Homo sapiens (Human).
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.

CX NCBI TaxID:9606;
 RN [1]

RP SEQUENCE FROM N.A.
 EX MEDLINE-92112994; PubMed-1730768,
 RA Kallunki P., Tytgavason K.;

RT "Human basement membrane heparan sulfate proteoglycan core protein: a
 RT 467-kD protein containing multiple domains resembling elements of the
 RT low density lipoprotein receptor, laminin, neural cell adhesion
 RT molecules, and epidermal growth factor.";
 RL J. Cell Biol. 116:559-571(1992).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE-Skin, and Colon;
 RX MEDLINE-92246394; PubMed-1569162;

RA Murdoch A.P., Dodge G.P., Cohen I., Tuan R.S., Iozzo R.V.;

RT "Primary structure of the human heparan sulfate proteoglycan from
 RT basement membrane (HSPG2/perlecan): A chimeric molecule with multiple
 RT domains homologous to the low density lipoprotein receptor, laminin,
 RT neural cell adhesion molecules, and epidermal growth factor.";
 RL J. Biol. Chem. 267:8544-8557(1992).
 RN [3]

RP SEQUENCE OF 1018-1472 FROM N.A.
 RC TISSUE-Colon;
 RX MEDLINE-91365376; PubMed-1679749;

RA Dodge G.R., Knvaitsky I., Chi M.L., Hassell J.P., McRidge G.W.,
 RA Yi H.F., Iozzo R.V.;

RT "Heparan sulfate proteoglycan of human colon: partial molecular
 RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
 RT short arm of human chromosome 1.";
 RL Genomics 10:673-680(1991).
 RN [4]

RP

RC SEQUENCE OF 892-1398 FROM N.A.
 RX TISSUE-Fibroblasts;
 EX MEDLINE-92110660; PubMed-1685141;

RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
 RA Tytgavason K.;

RT "Cloning of human heparan sulfate proteoglycan core protein,
 RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of
 RT a BamHI restriction fragment length polymorphism";
 RL Genomics 11:389-396(1991).
 RN [5]

RP SEQUENCE OF 1-21 FROM N.A.
 RX MEDLINE-94052171; PubMed-8234307;

RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;

RT "Structural characterization of the complete human perlecan gene and
 RT its promoter.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
 CC -1- FUNCTION: This protein is an integral component of basement
 CC membranes. It is responsible for the fixed negative electrostatic
 CC charge and is involved in the charge-selective ultrafiltration
 CC properties. It serves as an attachment substrate for cells.

CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.

CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.

CC -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
 CC AND O-LINKED OLIGOSACCHARIDES.

CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.

CC -1- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.

CC -1- SIMILARITY: CONTAINS 22 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: X62515; CAZ44373.1; -
 CC EMBL: M85289; AAS52700.1; -
 CC EMBL: M64283; AAS52699.1; -
 CC EMBL: S76436; AAB31121.2; -
 CC EMBL: L22078; -; NOT_ANNOTATED_CDS.
 CC HSPG: P00740; 1E0M.
 CC Slerna-2DPAGE; P9E160; -
 CC Genew; HGNC:5273; HSPG2.
 CC MIM; 142461; -
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001438; EGF_II.
 CC InterPro: IPR003406; Iq_MHC.
 CC InterPro: IPR003598; Iq_C2.
 CC InterPro: IPR002172; LDL_recept_A.
 CC InterPro: IPR000C34; Laminin_B.
 CC InterPro: IPR002649; Laminin_EGF.
 CC InterPro: IPR001791; Laminin_G.
 CC InterPro: IPR000C82; SEA_domain.
 CC Pfam: PF00047; Iq; 22.
 CC Pfam: PF00008; EGF; 4.
 CC Pfam: PF00052; laminin_B; 3.
 CC Pfam: PF00053; laminin_EGF; 7.
 CC Pfam: PF00054; laminin_G; 3.
 CC Pfam: PF00057; ldl_recept_a; 4.
 CC Pfam: PF01390; SEA; 1.
 CC PRINTS: PR00010; EGFBL00D.
 CC ProDom; PD003C31; Laminin_B; 3.
 CC SMART; SM0018C; EGF_Lam; 6.


```

QY 110 ELTANVFNK--VGLVWSESSYEAGTILSW--HLLGKPLVPNEKIVAVKEGIRKHPETSLFT 165
DB 158 RAKGKDKKVAIVACTISANKKDPUSVSWETHLKEG-----AEYGEIPNP-NGIVT 206
QY 166 LQSELMTTPARAGDPPTESNSFSLPHRALRTAPIGRVHVEVPLVEVLVVEP-EG 244
DB 207 VTSKRYLVPSK-EDHIFQSILACIVNVHMDREFRESLILNVQ---YEP-----EVTIESFDS 256
QY 225 GAVAFQCTVTHICFVPAQ-PSQIHWKDKGVLP---LPPSIVILIPICHCDOCTVSC 279
DB 257 NWTLRQMDVKLLKALANFPALEYHWTILNLSLPKAVEANNTLEFPAPINTSMASTYIP 316
QY 280 VATISSHIFESKA--VSTIIIE-----PREEPTASVW-----GSLPHTIALAUGILR 427
DB 317 EAIN-----PIGTHSQGVENITFEPTYPSPDPRGRRACQVPTAIIICGVVGSILILFVVG 372
QY 328 GAGTA 332
DB 373 GIVVA 377

RESULT 15
PGDB_MOUSE STANDARD: PRT: 3707 AA
ID PGDB_MOUSE
AC Q05793;
DF 01-NOV-1995 (Rel. 32, Created)
DF 01-NOV-1995 (Rel. 32, last sequence update)
DF 15-JUN-2002 (Rel. 41, last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PDC).
GN HSPG2.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N A
RC H153E-Melanoma;
RX MEDLINE=92078153; PubMed=1744097;
RA Noonan D.M., Horigan E.A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Hassell J.P.;
RT "The complete sequence of perlecan, a basement membrane heparan
RT sulfate proteoglycan, reveals extensive similarity with laminin A
RT chain, low density lipoprotein receptor, and the neural cell adhesion
RT molecule."
RL J. Biol. Chem. 266:22039-22047(1991)
RN [2]
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N A , AND PARTIAL SEQUENCE.
RX MEDLINE=89034110, PubMed=292709,
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogel C., Sasaki M.,
RA Yamada Y., Hassell J.P.;
RT "Identification of cDNA clones encoding different domains of the
RT basement membrane heparan sulfate proteoglycan."
RL J. Biol. Chem. 263:16379-16387(1988).
CC -1- FUNCTION: This protein is an integral component of basement
CC membranes. It is responsible for the fixed negative electrostatic
CC charge and is involved in the charge-selective ultrafiltration
CC properties. It serves as an attachment substrate for cells.
CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, proteoglycan and collagen type-
CC IV.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
CC -1- PTC: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED GLYCOSAMINOGLYCAN.
CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 10 5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 15 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

-1- SIMILARITY: CONTAINS 1 SEA DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: M77174; AAA39911.1; -
EMBL: J04954; AAA39991.1; -
EMBL: J04055; AAA39912.1; -
HSP: P01130; IAJJ.
MGD: MGI:96257; Hsp92.
InterPro: IPR000561; EGF-like.
InterPro: IPR000742; EGF_2.
InterPro: IPR001438; EGF_II.
InterPro: IPR003006; IG_MHC
InterPro: IPR003598; IG_C2.
InterPro: IPR002172; LDL_recept_A.
InterPro: IPR000034; Laminin_B.
InterPro: IPR002049; Laminin_EGF.
InterPro: IPR001791; Laminin_G.
InterPro: IPR000082; SEA_domain.
Pfam: PF00008; EGF; 4.
Pfam: PF00047; Ig; 15.
Pfam: PF00052; laminin_B; 3.
Pfam: PF00053; laminin_EGF; 8.
Pfam: PF00054; laminin_G; 3.
Pfam: PF00057; ldl_recept_a; 4.
Pfam: PF01390; SEA; 1.
PRINTS: PF00010; EGFRII00D.
Protein: P0003031; Laminin_B; 3.
SMART: SM00180; EGF_Like; 7.
SMART: SM00001; EGF_Like; 6.
SMART: SM00408; IGC2; 14.
SMART: SM00192; LDLa; 4.
SMART: SM00281; lamr; 3.
SMART: SM00282; lamc; 3.
SMART: SM00200; SEA; 1.
PROSITE: PS00022; EGF_1; 8.
PROSITE: PS01186; EGF_2; 5.
PROSITE: PS01248; LAMININ_TYPE_EGF; 11.
PROSITE: PS00025; LAMININ_DOMAIN_1; 4.
PROSITE: PS01209; LDLRA_1; 4.
PROSITE: PS00068; LDLRA_2; 4.
PROSITE: PS00024; SEA; 1.
Signal: Basement membrane; Proteoglycan; Repeat; Glycosylated;
KW Heparan sulfate, Laminin, EGF-like domain, Immunoglobulin domain,
KW Extracellular matrix, EGF-like domain.
FT SIGNAL: 1 21
FT CHAIN: 22 3707
FT DOMAIN: 80 194
FT DOMAIN: 195 234
FT DOMAIN: 281 319
FT DOMAIN: 320 359
FT DOMAIN: 360 403
FT DOMAIN: 404 504
FT DOMAIN: 521 530
FT DOMAIN: 531 730
FT DOMAIN: 731 763
FT DOMAIN: 764 813
FT DOMAIN: 814 871
FT DOMAIN: 879 923
FT DOMAIN: 924 933
FT DOMAIN: 934 1125
FT DOMAIN: 1126 1158
FT DOMAIN: 1159 1208
FT DOMAIN: 1209 1265
FT DOMAIN: 1275 1324
FT DOMAIN: 1325 1334
FT DOMAIN: 1335 1344
FT DOMAIN: 1345 1354
FT DOMAIN: 1355 1364
FT DOMAIN: 1365 1374
FT DOMAIN: 1375 1384
FT DOMAIN: 1385 1394
FT DOMAIN: 1395 1404
FT DOMAIN: 1405 1414
FT DOMAIN: 1415 1424
FT DOMAIN: 1425 1434
FT DOMAIN: 1435 1444
FT DOMAIN: 1445 1454
FT DOMAIN: 1455 1464
FT DOMAIN: 1465 1474
FT DOMAIN: 1475 1484
FT DOMAIN: 1485 1494
FT DOMAIN: 1495 1504
FT DOMAIN: 1505 1514
FT DOMAIN: 1515 1524
FT DOMAIN: 1525 1534
FT DOMAIN: 1535 1544
FT DOMAIN: 1545 1554
FT DOMAIN: 1555 1564
FT DOMAIN: 1565 1574
FT DOMAIN: 1575 1584
FT DOMAIN: 1585 1594
FT DOMAIN: 1595 1604
FT DOMAIN: 1605 1614
FT DOMAIN: 1615 1624
FT DOMAIN: 1625 1634
FT DOMAIN: 1635 1644
FT DOMAIN: 1645 1654
FT DOMAIN: 1655 1664
FT DOMAIN: 1665 1674
FT DOMAIN: 1675 1684
FT DOMAIN: 1685 1694
FT DOMAIN: 1695 1704
FT DOMAIN: 1705 1714
FT DOMAIN: 1715 1724
FT DOMAIN: 1725 1734
FT DOMAIN: 1735 1744
FT DOMAIN: 1745 1754
FT DOMAIN: 1755 1764
FT DOMAIN: 1765 1774
FT DOMAIN: 1775 1784
FT DOMAIN: 1785 1794
FT DOMAIN: 1795 1804
FT DOMAIN: 1805 1814
FT DOMAIN: 1815 1824
FT DOMAIN: 1825 1834
FT DOMAIN: 1835 1844
FT DOMAIN: 1845 1854
FT DOMAIN: 1855 1864
FT DOMAIN: 1865 1874
FT DOMAIN: 1875 1884
FT DOMAIN: 1885 1894
FT DOMAIN: 1895 1904
FT DOMAIN: 1905 1914
FT DOMAIN: 1915 1924
FT DOMAIN: 1925 1934
FT DOMAIN: 1935 1944
FT DOMAIN: 1945 1954
FT DOMAIN: 1955 1964
FT DOMAIN: 1965 1974
FT DOMAIN: 1975 1984
FT DOMAIN: 1985 1994
FT DOMAIN: 1995 2004
FT DOMAIN: 2005 2014
FT DOMAIN: 2015 2024
FT DOMAIN: 2025 2034
FT DOMAIN: 2035 2044
FT DOMAIN: 2045 2054
FT DOMAIN: 2055 2064
FT DOMAIN: 2065 2074
FT DOMAIN: 2075 2084
FT DOMAIN: 2085 2094
FT DOMAIN: 2095 2104
FT DOMAIN: 2105 2114
FT DOMAIN: 2115 2124
FT DOMAIN: 2125 2134
FT DOMAIN: 2135 2144
FT DOMAIN: 2145 2154
FT DOMAIN: 2155 2164
FT DOMAIN: 2165 2174
FT DOMAIN: 2175 2184
FT DOMAIN: 2185 2194
FT DOMAIN: 2195 2204
FT DOMAIN: 2205 2214
FT DOMAIN: 2215 2224
FT DOMAIN: 2225 2234
FT DOMAIN: 2235 2244
FT DOMAIN: 2245 2254
FT DOMAIN: 2255 2264
FT DOMAIN: 2265 2274
FT DOMAIN: 2275 2284
FT DOMAIN: 2285 2294
FT DOMAIN: 2295 2304
FT DOMAIN: 2305 2314
FT DOMAIN: 2315 2324
FT DOMAIN: 2325 2334
FT DOMAIN: 2335 2344
FT DOMAIN: 2345 2354
FT DOMAIN: 2355 2364
FT DOMAIN: 2365 2374
FT DOMAIN: 2375 2384
FT DOMAIN: 2385 2394
FT DOMAIN: 2395 2404
FT DOMAIN: 2405 2414
FT DOMAIN: 2415 2424
FT DOMAIN: 2425 2434
FT DOMAIN: 2435 2444
FT DOMAIN: 2445 2454
FT DOMAIN: 2455 2464
FT DOMAIN: 2465 2474
FT DOMAIN: 2475 2484
FT DOMAIN: 2485 2494
FT DOMAIN: 2495 2504
FT DOMAIN: 2505 2514
FT DOMAIN: 2515 2524
FT DOMAIN: 2525 2534
FT DOMAIN: 2535 2544
FT DOMAIN: 2545 2554
FT DOMAIN: 2555 2564
FT DOMAIN: 2565 2574
FT DOMAIN: 2575 2584
FT DOMAIN: 2585 2594
FT DOMAIN: 2595 2604
FT DOMAIN: 2605 2614
FT DOMAIN: 2615 2624
FT DOMAIN: 2625 2634
FT DOMAIN: 2635 2644
FT DOMAIN: 2645 2654
FT DOMAIN: 2655 2664
FT DOMAIN: 2665 2674
FT DOMAIN: 2675 2684
FT DOMAIN: 2685 2694
FT DOMAIN: 2695 2704
FT DOMAIN: 2705 2714
FT DOMAIN: 2715 2724
FT DOMAIN: 2725 2734
FT DOMAIN: 2735 2744
FT DOMAIN: 2745 2754
FT DOMAIN: 2755 2764
FT DOMAIN: 2765 2774
FT DOMAIN: 2775 2784
FT DOMAIN: 2785 2794
FT DOMAIN: 2795 2804
FT DOMAIN: 2805 2814
FT DOMAIN: 2815 2824
FT DOMAIN: 2825 2834
FT DOMAIN: 2835 2844
FT DOMAIN: 2845 2854
FT DOMAIN: 2855 2864
FT DOMAIN: 2865 2874
FT DOMAIN: 2875 2884
FT DOMAIN: 2885 2894
FT DOMAIN: 2895 2904
FT DOMAIN: 2905 2914
FT DOMAIN: 2915 2924
FT DOMAIN: 2925 2934
FT DOMAIN: 2935 2944
FT DOMAIN: 2945 2954
FT DOMAIN: 2955 2964
FT DOMAIN: 2965 2974
FT DOMAIN: 2975 2984
FT DOMAIN: 2985 2994
FT DOMAIN: 2995 3004
FT DOMAIN: 3005 3014
FT DOMAIN: 3015 3024
FT DOMAIN: 3025 3034
FT DOMAIN: 3035 3044
FT DOMAIN: 3045 3054
FT DOMAIN: 3055 3064
FT DOMAIN: 3065 3074
FT DOMAIN: 3075 3084
FT DOMAIN: 3085 3094
FT DOMAIN: 3095 3104
FT DOMAIN: 3105 3114
FT DOMAIN: 3115 3124
FT DOMAIN: 3125 3134
FT DOMAIN: 3135 3144
FT DOMAIN: 3145 3154
FT DOMAIN: 3155 3164
FT DOMAIN: 3165 3174
FT DOMAIN: 3175 3184
FT DOMAIN: 3185 3194
FT DOMAIN: 3195 3204
FT DOMAIN: 3205 3214
FT DOMAIN: 3215 3224
FT DOMAIN: 3225 3234
FT DOMAIN: 3235 3244
FT DOMAIN: 3245 3254
FT DOMAIN: 3255 3264
FT DOMAIN: 3265 3274
FT DOMAIN: 3275 3284
FT DOMAIN: 3285 3294
FT DOMAIN: 3295 3304
FT DOMAIN: 3305 3314
FT DOMAIN: 3315 3324
FT DOMAIN: 3325 3334
FT DOMAIN: 3335 3344
FT DOMAIN: 3345 3354
FT DOMAIN: 3355 3364
FT DOMAIN: 3365 3374
FT DOMAIN: 3375 3384
FT DOMAIN: 3385 3394
FT DOMAIN: 3395 3404
FT DOMAIN: 3405 3414
FT DOMAIN: 3415 3424
FT DOMAIN: 3425 3434
FT DOMAIN: 3435 3444
FT DOMAIN: 3445 3454
FT DOMAIN: 3455 3464
FT DOMAIN: 3465 3474
FT DOMAIN: 3475 3484
FT DOMAIN: 3485 3494
FT DOMAIN: 3495 3504
FT DOMAIN: 3505 3514
FT DOMAIN: 3515 3524
FT DOMAIN: 3525 3534
FT DOMAIN: 3535 3544
FT DOMAIN: 3545 3554
FT DOMAIN: 3555 3564
FT DOMAIN: 3565 3574
FT DOMAIN: 3575 3584
FT DOMAIN: 3585 3594
FT DOMAIN: 3595 3604
FT DOMAIN: 3605 3614
FT DOMAIN: 3615 3624
FT DOMAIN: 3625 3634
FT DOMAIN: 3635 3644
FT DOMAIN: 3645 3654
FT DOMAIN: 3655 3664
FT DOMAIN: 3665 3674
FT DOMAIN: 3675 3684
FT DOMAIN: 3685 3694
FT DOMAIN: 3695 3704
FT DOMAIN: 3705 3714
FT DOMAIN: 3715 3724
FT DOMAIN: 3725 3734
FT DOMAIN: 3735 3744
FT DOMAIN: 3745 3754
FT DOMAIN: 3755 3764
FT DOMAIN: 3765 3774
FT DOMAIN: 3775 3784
FT DOMAIN: 3785 3794
FT DOMAIN: 3795 3804
FT DOMAIN: 3805 3814
FT DOMAIN: 3815 3824
FT DOMAIN: 3825 3834
FT DOMAIN: 3835 3844
FT DOMAIN: 3845 3854
FT DOMAIN: 3855 3864
FT DOMAIN: 3865 3874
FT DOMAIN: 3875 3884
FT DOMAIN: 3885 3894
FT DOMAIN: 3895 3904
FT DOMAIN: 3905 3914
FT DOMAIN: 3915 3924
FT DOMAIN: 3925 3934
FT DOMAIN: 3935 3944
FT DOMAIN: 3945 3954
FT DOMAIN: 3955 3964
FT DOMAIN: 3965 3974
FT DOMAIN: 3975 3984
FT DOMAIN: 3985 3994
FT DOMAIN: 3995 4004
FT DOMAIN: 4005 4014
FT DOMAIN: 4015 4024
FT DOMAIN: 4025 4034
FT DOMAIN: 4035 4044
FT DOMAIN: 4045 4054
FT DOMAIN: 4055 4064
FT DOMAIN: 4065 4074
FT DOMAIN: 4075 4084
FT DOMAIN: 4085 4094
FT DOMAIN: 4095 4104
FT DOMAIN: 4105 4114
FT DOMAIN: 4115 4124
FT DOMAIN: 4125 4134
FT DOMAIN: 4135 4144
FT DOMAIN: 4145 4154
FT DOMAIN: 4155 4164
FT DOMAIN: 4165 4174
FT DOMAIN: 4175 4184
FT DOMAIN: 4185 4194
FT DOMAIN: 4195 4204
FT DOMAIN: 4205 4214
FT DOMAIN: 4215 4224
FT DOMAIN: 4225 4234
FT DOMAIN: 4235 4244
FT DOMAIN: 4245 4254
FT DOMAIN: 4255 4264
FT DOMAIN: 4265 4274
FT DOMAIN: 4275 4284
FT DOMAIN: 4285 4294
FT DOMAIN: 4295 4304
FT DOMAIN: 4305 4314
FT DOMAIN: 4315 4324
FT DOMAIN: 4325 4334
FT DOMAIN: 4335 4344
FT DOMAIN: 4345 4354
FT DOMAIN: 4355 4364
FT DOMAIN: 4365 4374
FT DOMAIN: 4375 4384
FT DOMAIN: 4385 4394
FT DOMAIN: 4395 4404
FT DOMAIN: 4405 4414
FT DOMAIN: 4415 4424
FT DOMAIN: 4425 4434
FT DOMAIN: 4435 4444
FT DOMAIN: 4445 4454
FT DOMAIN: 4455 4464
FT DOMAIN: 4465 4474
FT DOMAIN: 4475 4484
FT DOMAIN: 4485 4494
FT DOMAIN: 4495 4504
FT DOMAIN: 4505 4514
FT DOMAIN: 4515 4524
FT DOMAIN: 4525 4534
FT DOMAIN: 4535 4544
FT DOMAIN: 4545 4554
FT DOMAIN: 4555 4564
FT DOMAIN: 4565 4574
FT DOMAIN: 4575 4584
FT DOMAIN: 4585 4594
FT DOMAIN: 4595 4604
FT DOMAIN: 4605 4614
FT DOMAIN: 4615 4624
FT DOMAIN: 4625 4634
FT DOMAIN: 4635 4644
FT DOMAIN: 4645 4654
FT DOMAIN: 4655 4664
FT DOMAIN: 4665 4674
FT DOMAIN: 4675 4684
FT DOMAIN: 4685 4694
FT DOMAIN: 4695 4704
FT DOMAIN: 4705 4714
FT DOMAIN: 4715 4724
FT DOMAIN: 4725 4734
FT DOMAIN: 4735 4744
FT DOMAIN: 4745 4754
FT DOMAIN: 4755 4764
FT DOMAIN: 4765 4774
FT DOMAIN: 4775 4784
FT DOMAIN: 4785 4794
FT DOMAIN: 4795 4804
FT DOMAIN: 4805 4814
FT DOMAIN: 4815 4824
FT DOMAIN: 4825 4834
FT DOMAIN: 4835 4844
FT DOMAIN: 4845 4854
FT DOMAIN: 4855 4864
FT DOMAIN: 4865 4874
FT DOMAIN: 4875 4884
FT DOMAIN: 4885 4894
FT DOMAIN: 4895 4904
FT DOMAIN: 4905 4914
FT DOMAIN: 4915 4924
FT DOMAIN: 4925 4934
FT DOMAIN: 4935 4944
FT DOMAIN: 4945 4954
FT DOMAIN: 4955 4964
FT DOMAIN: 4965 4974
FT DOMAIN: 4975 4984
FT DOMAIN: 4985 4994
FT DOMAIN: 4995 5004
FT DOMAIN: 5005 5014
FT DOMAIN: 5015 5024
FT DOMAIN: 5025 5034
FT DOMAIN: 5035 5044
FT DOMAIN: 5045 5054
FT DOMAIN: 5055 5064
FT DOMAIN: 5065 5074
FT DOMAIN: 5075 5084
FT DOMAIN: 5085 5094
FT DOMAIN: 5095 5104
FT DOMAIN: 5105 5114
FT DOMAIN: 5115 5124
FT DOMAIN: 5125 5134
FT DOMAIN: 5135 5144
FT DOMAIN: 5145 5154
FT DOMAIN: 5155 5164
FT DOMAIN: 5165 5174
FT DOMAIN: 5175 5184
FT DOMAIN: 5185 5194
FT DOMAIN: 5195 5204
FT DOMAIN: 5205 5214
FT DOMAIN: 5215 5224
FT DOMAIN: 5225 5234
FT DOMAIN: 5235 5244
FT DOMAIN: 5245 5254
FT DOMAIN: 5255 5264
FT DOMAIN: 5265 5274
FT DOMAIN: 5275 5284
FT DOMAIN: 5285 5294
FT DOMAIN: 5295 5304
FT DOMAIN: 5305 5314
FT DOMAIN: 5315 5324
FT DOMAIN: 5325 5334
FT DOMAIN: 5335 5344
FT DOMAIN: 5345 5354
FT DOMAIN: 5355 5364
FT DOMAIN: 5365 5374
FT DOMAIN: 5375 5384
FT DOMAIN: 5385 5394
FT DOMAIN: 5395 5404
FT DOMAIN: 5405 5414
FT DOMAIN: 5415 5424
FT DOMAIN: 5425 5434
FT DOMAIN: 5435 5444
FT DOMAIN: 5445 5454
FT DOMAIN: 5455 5464
FT DOMAIN: 5465 5474
FT DOMAIN: 5475 5484
FT DOMAIN: 5485 5494
FT DOMAIN: 5495 5504
FT DOMAIN: 5505 5514
FT DOMAIN: 5515 5524
FT DOMAIN: 5525 5534
FT DOMAIN: 5535 5544
FT DOMAIN: 5545 5554
FT DOMAIN: 5555 5564
FT DOMAIN: 5565 5574
FT DOMAIN: 5575 5584
FT DOMAIN: 5585 5594
FT DOMAIN: 5595 5604
FT DOMAIN: 5605 5614
FT DOMAIN: 5615 5624
FT DOMAIN: 5625 5634
FT DOMAIN: 5635 5644
FT DOMAIN: 5645 5654
FT DOMAIN: 5655 5664
FT DOMAIN: 5665 5674
FT DOMAIN: 5675 5684
FT DOMAIN: 5685 5694
FT DOMAIN: 5695 5704
FT DOMAIN: 5705 5714
FT DOMAIN: 5715 5724
FT DOMAIN: 5725 5734
FT DOMAIN: 5735 5744
FT DOMAIN: 5745 5754
FT DOMAIN: 5755 5764
FT DOMAIN: 5765 5774
FT DOMAIN: 5775 5784
FT DOMAIN: 5785 5794
FT DOMAIN: 5795 5804
FT DOMAIN: 5805 5814
FT DOMAIN: 5815 5824
FT DOMAIN: 5825 5834
FT DOMAIN: 5835 5844
FT DOMAIN: 5845 5854
FT DOMAIN: 5855 5864
FT DOMAIN: 5865 5874
FT DOMAIN: 5875 5884
FT DOMAIN: 5885 5894
FT DOMAIN: 5895 5904
FT DOMAIN: 5905 5914
FT DOMAIN: 5915 5924
FT DOMAIN: 5925 5934
FT DOMAIN: 5935 5944
FT DOMAIN: 5945 5954
FT DOMAIN: 5955 5964
FT DOMAIN: 5965 5974
FT DOMAIN: 5975 5984
FT DOMAIN: 5985 5994
FT DOMAIN: 5995 6004
FT DOMAIN: 6005 6014
FT DOMAIN: 6015 6024
FT DOMAIN: 6025 6034
FT DOMAIN: 6035 6044
FT DOMAIN: 6045 6054
FT DOMAIN: 6055 6064
FT DOMAIN: 6065 6074
FT DOMAIN: 6075 6084
FT DOMAIN: 6085 6094
FT DOMAIN: 6095 6104
FT DOMAIN: 6105 6114
FT DOMAIN: 6115 6124
FT DOMAIN: 6125 6134
FT DOMAIN: 6135 6144
FT DOMAIN: 6145 6154
FT DOMAIN: 6155 6164
FT DOMAIN: 6165 6174
FT DOMAIN: 6175 6184
FT DOMAIN: 6185 6194
FT DOMAIN: 6195 6204
FT DOMAIN: 6205 6214
FT DOMAIN: 6215 6224
FT DOMAIN: 6225 6234
FT DOMAIN: 6235 6244
FT DOMAIN: 6245 6254
FT DOMAIN: 6255 6264
FT DOMAIN: 6265 6274
FT DOMAIN: 6275 6284
FT DOMAIN: 6285 6294
FT DOMAIN: 6295 6304
FT DOMAIN: 6305 6314
FT DOMAIN: 6315 6324
FT DOMAIN: 6325 6334
FT DOMAIN: 6335 6344
FT DOMAIN: 6345 6354
FT DOMAIN: 6355 6364
FT DOMAIN: 6365 6374
FT DOMAIN: 6375 6384
FT DOMAIN: 6385 6394
FT DOMAIN: 6395 6404
FT DOMAIN: 6405 6414
FT DOMAIN: 6415 6424
FT DOMAIN: 6425 6434
FT DOMAIN: 6435 6444
FT DOMAIN: 6445 6454
FT DOMAIN: 6455 6464
FT DOMAIN: 6465 6474
FT DOMAIN: 6475 6484
FT DOMAIN: 6485 6494
FT DOMAIN: 6495 6504
FT DOMAIN: 6505 6514
FT DOMAIN: 6515 6524
FT DOMAIN: 6525 6534
FT DOMAIN: 6535 6544
FT DOMAIN: 6545 6554
FT DOMAIN: 6555 6564
FT DOMAIN: 6565 6574
FT DOMAIN: 6575 6584
FT DOMAIN: 6585 6594
FT DOMAIN: 6595 6604
FT DOMAIN: 6605 6614
FT DOMAIN: 6615 6624
FT DOMAIN: 6625 6634
FT DOMAIN: 6635 6644
FT DOMAIN: 6645 6654
FT DOMAIN: 6655 6664
FT DOMAIN: 6665 6674
FT DOMAIN: 6675 6684
FT DOMAIN: 6685 6694
FT DOMAIN: 6695 6704
FT DOMAIN: 6705 6714
FT DOMAIN: 6715 6724
FT DOMAIN: 6725 6734
FT DOMAIN: 6735 6744
FT DOMAIN: 6745 6754
FT DOMAIN: 6755 6764
FT DOMAIN: 6765 6774
FT DOMAIN: 6775 6784
FT DOMAIN: 6785 6794
FT DOMAIN: 6795 6804
FT DOMAIN: 6805 6814
FT DOMAIN: 6815 6824
FT DOMAIN: 6825 6834
FT DOMAIN: 6835 6844
FT DOMAIN: 6845 6854
FT DOMAIN: 6855 6864
FT DOMAIN: 6865 6874
FT DOMAIN: 6875 6884
FT DOMAIN: 6885 6894
FT DOMAIN: 6895 6904
FT DOMAIN: 6905 6914
FT DOMAIN: 6915 6924
FT DOMAIN: 6925 6934
FT DOMAIN: 6935 6944
FT DOMAIN: 6945 6954
FT DOMAIN: 6955 6964
FT DOMAIN: 6965 6974
FT DOMAIN: 6975 6984
FT DOMAIN: 6985 6994
FT DOMAIN: 6995 7004
FT DOMAIN: 7005 7014
FT DOMAIN: 7015 7024
FT DOMAIN: 7025 7034
FT DOMAIN: 7035 7044
FT DOMAIN: 7045 7054
FT DOMAIN: 7055 7064
FT DOMAIN: 7065 7074
FT DOMAIN: 7075 7084
FT DOMAIN: 7085 7094
FT DOMAIN: 7095 7104
FT DOMAIN: 7105 7114
FT DOMAIN: 7115 7124
FT DOMAIN: 7125 7134
FT DOMAIN: 7135 7144
FT DOMAIN: 7145 7154
FT DOMAIN: 7155 7164
FT DOMAIN: 7165 7174
FT DOMAIN: 7175 7184
FT DOMAIN: 7185 7194
FT DOMAIN: 7195 7204
FT DOMAIN: 7205 7214
FT DOMAIN: 7215 7224
FT DOMAIN: 7225 7234
FT DOMAIN: 7235 7244
FT DOMAIN: 7245 7254
FT DOMAIN: 7255 7264
FT DOMAIN: 7265 7274
FT DOMAIN: 7275 7284
FT DOMAIN: 7285 7294
FT DOMAIN: 7295 7304
FT DOMAIN: 7305 7314
FT DOMAIN: 7315 7324
FT DOMAIN: 7325 7334
FT DOMAIN: 7335 7344
FT DOMAIN: 7345 7354
FT DOMAIN: 7355 7364
FT DOMAIN: 7365 7374
FT DOMAIN: 7375 7384
FT DOMAIN: 7385 7394
FT DOMAIN: 7395 7404
FT DOMAIN: 7405 7414
FT DOMAIN: 7415 7424
FT DOMAIN: 7425 7434
FT DOMAIN: 7435 7444
FT DOMAIN: 7445 7454
FT DOMAIN: 7455 7464
FT DOMAIN: 7465 7474
FT DOMAIN: 7475 7484
FT DOMAIN: 7485 7494
FT DOMAIN: 7495 7504
FT DOMAIN: 7505 7514
FT DOMAIN: 7515 7524
FT DOMAIN: 7525 7534
FT DOMAIN: 7535 7544
FT DOMAIN: 7545 7554
FT DOMAIN: 7555 7564
FT DOMAIN: 7565 7574
FT DOMAIN: 7575 7584
FT DOMAIN: 7585 7594
FT DOMAIN: 7595 7604
FT DOMAIN: 7605 7614
FT DOMAIN: 7615 7624
FT DOMAIN: 7625 7634
FT DOMAIN: 7635 7644
FT DOMAIN: 7645 7654
FT DOMAIN: 7655 7664
FT DOMAIN: 7665 7674
FT DOMAIN: 7675 7684
FT DOMAIN: 7685 7694
FT DOMAIN: 7695 7704
FT DOMAIN: 7705 7714
FT DOMAIN: 7715 7724
FT DOMAIN: 7725 7734
FT DOMAIN: 7735 7744
FT DOMAIN: 7745 7754
FT DOMAIN: 7755 7764
FT DOMAIN: 7765 7774
FT DOMAIN: 7775 7784
FT DOMAIN: 7785 7794
FT DOMAIN: 7795 7804
FT DOMAIN: 7805 7814
FT DOMAIN: 7815 7824
FT DOMAIN: 7825 7834
FT DOMAIN: 7835 7844
FT DOMAIN: 7845 7854
FT DOMAIN: 7855 7864
FT DOMAIN: 7865 7874
FT DOMAIN: 7875 7884
FT DOMAIN: 7885 7894
FT DOMAIN: 7895 7904
FT DOMAIN: 7905 7914
FT DOMAIN: 7915 7924
FT DOMAIN: 7925 7934
FT DOMAIN: 7935 7944
FT DOMAIN: 7945 7954
FT DOMAIN: 7955 7964
FT DOMAIN: 7965 7974
FT DOMAIN: 7975 7984
FT DOMAIN: 7985 7994
FT DOMAIN: 7995 8004
FT DOMAIN: 8005 8014
FT DOMAIN: 8015 8024
FT DOMAIN: 8025 8034
FT DOMAIN: 8035 8044
FT DOMAIN: 8045 8054
FT DOMAIN: 8055 8064
FT DOMAIN: 8065 8074
FT DOMAIN: 8075 8084
FT DOMAIN: 8085 8094
FT DOMAIN: 8095 8104
FT DOMAIN: 8105 8114
FT DOMAIN: 8115 8124
FT DOMAIN: 8125 8134
FT DOMAIN: 8135 8144
FT DOMAIN: 8145 8154
FT DOMAIN: 8155 8164
FT DOMAIN: 8165 8174
FT DOMAIN: 8175 8184
FT DOMAIN: 8185 8194
FT DOMAIN: 8195 8204
FT DOMAIN: 8205 8214
FT DOMAIN: 8215 8224
FT DOMAIN: 8225 8234
FT DOMAIN: 8235 8244
FT DOMAIN: 8245 8254
FT DOMAIN: 8255 8264
FT DOMAIN: 8265 8274
FT DOMAIN: 8275 8284
FT DOMAIN: 8285 8294
FT DOMAIN: 8295 8304
FT DOMAIN: 8305 8314
FT DOMAIN: 8315 8324
FT DOMAIN: 8325 8334
FT DOMAIN: 8335 8344
FT DOMAIN: 8345 8354
FT DOMAIN: 8355 8364
FT DOMAIN: 8365 837
```

FL	DOMAIN	1335	1529	AMININ DOMAIN IV 3 (DOMAIN 111 C)
FT	DOMAIN	1530	1562	AMININ EGF-LIKE 9 (C TERMINAL)
FT	DOMAIN	1563	1612	AMININ EGF-LIKE 10
FT	DOMAIN	1613	1670	AMININ EGF-LIKE 11
FT	DOMAIN	1671	1771	IG LIKE C2 TYPE DOMAIN 2
FL	DOMAIN	1772	1865	IG LIKE C2 TYPE DOMAIN 4
FL	DOMAIN	1866	1954	IG LIKE C2 TYPE DOMAIN 4
FT	DOMAIN	1955	1999	IG LIKE C2 TYPE DOMAIN 5
FT	DOMAIN	2050	2148	IG LIKE C2 TYPE DOMAIN 6
FT	DOMAIN	2149	2244	IG LIKE C2 TYPE DOMAIN 7
FT	DOMAIN	2245	2343	IG LIKE C2 TYPE DOMAIN 8
FL	DOMAIN	2344	2436	IG LIKE C2 TYPE DOMAIN 9
FL	DOMAIN	2437	2532	IG LIKE C2 TYPE DOMAIN 10
FT	DOMAIN	2533	2619	IG LIKE C2 TYPE DOMAIN 11
FT	DOMAIN	2620	2720	IG LIKE C2 TYPE DOMAIN 12
FL	DOMAIN	2721	2809	IG LIKE C2 TYPE DOMAIN 13
FL	DOMAIN	2810	2896	IG LIKE C2 TYPE DOMAIN 14
FL	DOMAIN	2896	2980	IG LIKE C2 TYPE DOMAIN 15
FL	DOMAIN	2984	3162	IG LIKE C2 TYPE DOMAIN 15
FL	DOMAIN	3163	3241	EGF-LIKE
FL	DOMAIN	3245	3425	AMININ G-LIKE 2
FT	DOMAIN	3518	3705	AMININ G-LIKE 3
FT	SITE	65	67	HEPARAN SULFATE (POTENTIAL)
FL	SITE	71	73	HEPARAN SULFATE (POTENTIAL)
FL	SITE	76	78	HEPARAN SULFATE (POTENTIAL)
FL	SITE	3615	3617	MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL)
FT	DISULFID	199	212	BY SIMILARITY
FT	DISULFID	206	225	BY SIMILARITY
FL	DISULFID	219	234	BY SIMILARITY
FL	DISULFID	285	297	BY SIMILARITY
FL	DISULFID	292	310	BY SIMILARITY
FL	DISULFID	404	319	BY SIMILARITY
FL	DISULFID	425	337	BY SIMILARITY
FT	DISULFID	432	450	BY SIMILARITY
FL	DISULFID	344	359	BY SIMILARITY
FL	DISULFID	368	381	BY SIMILARITY
FL	DISULFID	375	394	BY SIMILARITY
FL	DISULFID	488	403	BY SIMILARITY
FL	DISULFID	498	479	BY SIMILARITY
FL	DISULFID	764	773	BY SIMILARITY
FL	DISULFID	766	780	BY SIMILARITY
FL	DISULFID	783	792	BY SIMILARITY
FL	DISULFID	795	811	BY SIMILARITY
FT	DISULFID	814	829	BY SIMILARITY
FL	DISULFID	816	839	BY SIMILARITY
FL	DISULFID	842	851	BY SIMILARITY
FL	DISULFID	854	869	BY SIMILARITY
FL	DISULFID	1159	1168	BY SIMILARITY
FL	DISULFID	1161	1175	BY SIMILARITY
FL	DISULFID	1178	1187	BY SIMILARITY
FL	DISULFID	1190	1206	BY SIMILARITY
FT	DISULFID	1209	1224	BY SIMILARITY
FL	DISULFID	1211	1234	BY SIMILARITY
FL	DISULFID	1247	1246	BY SIMILARITY
FL	DISULFID	1249	1263	BY SIMILARITY
FL	DISULFID	1275	1287	BY SIMILARITY
FT	DISULFID	1277	1293	BY SIMILARITY
FT	DISULFID	1295	1304	BY SIMILARITY
FL	DISULFID	1307	1322	BY SIMILARITY
FL	DISULFID	1563	1572	BY SIMILARITY
FL	DISULFID	1565	1579	BY SIMILARITY
FT	DISULFID	1582	1591	BY SIMILARITY
FL	DISULFID	1594	1610	BY SIMILARITY
FL	DISULFID	1613	1628	BY SIMILARITY
FL	DISULFID	1615	1638	BY SIMILARITY
FL	DISULFID	1641	1650	BY SIMILARITY
FL	DISULFID	1653	1668	BY SIMILARITY
FL	DISULFID	1792	1809	BY SIMILARITY
FT	DISULFID	1886	1932	BY SIMILARITY
FL	DISULFID	1976	2021	BY SIMILARITY
FL	DISULFID	2073	2118	BY SIMILARITY
FL	DISULFID	2170	2215	BY SIMILARITY

[illegible]

Search completed: May 30, 2003, 15:55:49
Job time : 21.4414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search using sw model

Run On: May 30, 2003, 16:47:46; Search time 77.018 seconds
(without alignments)
888,203 Million cell updates/sec

Title: US-09-872-185B-2
Percent score: 1766
Sequence: 1 AUNTHARIGELPLVAKKAGAP.....ESSELTALALAHIDHDATA 332

Scoring table: BLQSUM62
Gap 10.0, Gapext 0.5

Searched: 671580 seqs, 20647115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 255555555

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL-21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriopl.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1381	78.2	402	11	035444
2	224.5	12.7	583	11	Q8R2T0
3	220	12.5	521	6	Q46651
4	219.5	12.4	646	4	Q95812
5	218.5	12.4	646	4	Q95812
6	217	12.3	1344	11	Q95234
7	214	12.1	1341	4	Q95234
8	213	12.1	583	11	Q35112
9	210.5	11.9	340	13	Q95243
10	210	11.9	1496	4	Q95234
11	209	11.8	1612	11	Q95234
12	209	11.8	1651	4	Q95234
13	208	11.8	1043	5	Q24327
14	204.5	11.6	583	6	Q95813
15	204	11.6	521	6	Q46634
16	204	11.6	5636	4	Q95817

17	203	11.5	1033	5	Q9V643
18	203	11.5	1651	11	Q55005
19	201	11.4	743	11	Q70246
20	201	11.4	1389	13	Q90289
21	201	11.4	1419	13	Q98SW3
22	198.5	11.3	606	11	Q9ESS8
23	198.5	11.2	548	11	Q9EF92
24	198	11.2	4770	4	Q9H3V5
25	196.5	11.1	1614	13	Q90V07
26	196	11.1	333	13	Q95241
27	195.5	11.1	1380	4	Q9HCK4
28	195	11.0	1102	11	Q923W7
29	194.5	11.0	584	13	Q98Q21
30	194.5	11.0	584	13	Q90989
31	194.5	11.0	584	13	Q98Q21
32	194.5	11.0	626	13	Q98Q21
33	194.5	11.0	626	13	Q98Q21
34	192.5	10.9	1032	13	Q95VD6
35	192.5	10.9	1342	11	Q9EQS8
36	192.5	10.9	1253	11	Q9EQS8
37	191	10.8	1675	13	Q98SW4
38	189.5	10.7	1273	5	Q44928
39	185.5	10.6	1513	13	Q90Z70
40	185.5	10.6	444	4	Q94121
41	185.5	10.5	458	11	Q61351
42	185.5	10.5	521	11	Q61352
43	185.5	10.5	939	5	Q967X6
44	184.5	10.4	444	11	Q95P00
45	183.5	10.4	1114	4	Q9BWL1

ALIGNMENTS

RESULT 1

035444

ID: 035444 PRELIMINARY: PRT: 402 AA.

AC 035444; (TREMBL: 05, Created)

DT 01-JAN-1998 (TREMBL: 05, Last sequence update)

DT 01-DEC-2001 (TREMBL: 19, Last annotation update)

DE RAGE.

GN RAGE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

OX NCBI_FaxID=10090;

RN [1]

PP SEQUENCE FROM N.A.

PA Bowen L, Mahairas G, Qin S, Ahearn M.E., Dankers C., Lasky S.,

KA Lorrain C, Schmitt S, Tipton S, Traicoff R, Zaakore K, Hood L;

RT "Sequence of the mouse major histocompatibility locus class III

RT region."

RL Submitted (01-1997) to the EMBL/GenBank/DBJ databases.

DR EMRL: AF010001; ZABR2007 1;

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_Like.

DR InterPro: IPR003606; Ig_MHC.

DR Pfam: PF00047; Ig; 3.

DR SMART: SM00408, IGC2; 1.

DR SMART: SM00416, IG_Like; 1.

DR PROSITE: PS00346, IG_MHC; UNKNOWN_1.

KW Immunoglobulin domain.

SU SUGGESTER: 402 AA; 4255 MW; DBF025A0C0B992 CRC64;

Query Match: 78.2%, Score 1381, DB 11, Length 402.

Best local similarity: 79.9%, Pred. No. 5, 100.

Matches: 263, Conservative: 19, Mismatches: 45, Indels: 2, Gaps: 2.

Qy 2 QNIAKIGHIVLKKCAIKKKIKVKKLNTGTEANKVLSPQGGWDSVARVLPNGS 61

24 QNITAPTEPIVLSPKFAKKKKQQLKWLKNTGTEANKVLSPQGGWDSVARVLPNGS 82

Db 92 LGDCISRDSPEGNITWYRNGKVLQPLEGAVIIFKKQMDPVLTQTYMTSSLEYKTK-A 150
 QY 179 DPRPTFSCSPGIPRRKALRTAPIQPRVWE-DVPLEEVLVVEPEGCAVAPGTVTLQC 237
 Db 151 DQGFPPGSDIYYGCSQ--KVSHPGAVFDIYYPGQVTLQVLPKNAIKEDUNITLKC 208
 QY 238 EVPAQSP-QIHWKDKVPLPLPSPVLLIPEIGPDQGYSC 279
 Db 209 LQCNPTTFPPFVILGCPGCISSNTYTLNVRNRATGNYKC 251

RESULT 4
 Q95812
 ID Q95812 PRELIMINARY; PRT; 646 AA.
 AC Q95812;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Cell surface glycoprotein PHLI2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 PP SEQUENCE FROM N.A.
 RA Gui L., Chang L., Browne P.V., Heibel R.P.;
 RT "PHI2 from human umbilical vein endothelial cells";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF084868; AAL7794.1;
 DR InterPro: IPR003594; Ig_c2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 5.
 DR SMART: SM00408; IGc2; 2.
 DR SMART: SM00410; IG_Like; 2.
 KW Immunoglobulin domain.
 SQ SEQUENCE 646 AA; 71665 MW; 1B5FB8ADD930738E CRC64;

Query Match 12.48; Score 219.5; DB 4; Length 646;
 Best Local Similarity 23.38; Pred. No. 6e-09;
 Matches 95; Conservative 56; Mismatches 134; Indels 123; Gaps 19;

QY 4 ITARIGELPLVKYKGAPKKPP-QRLW--KLNTGRTEAWKVLSPG---GGPWDSVARVL 57
 Db 36 VEVEVGSTALLKCGLSQSGNLSHVDFSVHKKRKTLLIFRVGQGGQSEGEYEQRLSIQ 95
 QY 58 PNG-SLFLPAGVIGDEGIFRCQAMNRNKETKS-NYRV--RYVOIPGKPEI-----V 105
 Db 96 DRGATIALTVLPQDERIFLQ-----GKPPKQVEYRIQLPVYKAPPEPNQVNPGLGIPV 150
 QY 106 DSASELTAGVNVKVTGCVSGSYDAGTILSWHLDGKPLVPNEKGVSVKQTRRHPTGLFT 165
 Db 151 NSKF-----PPFVATCVGNGYPIQVIVYKNGRPLKEKNRVHI-QSSQTVSESSGLYT 203
 QY 166 LQSELMTVPARKSDPPPTFS--SFSP-ILPHFALP-----TAPI---QPPVW---EPVPL- 213
 Db 204 LQSLILKAQLVK-EDKDAQFYCELNYRLPSGNHMKESREVTVPFYPTKVMLEVEPVGML 262
 QY 214 -----EE----- 215
 Db 263 KEDRVEIKCLADGNPFPFHSISEUNPSTREAEETTNDUNGVLVLEPARKHESRYEQG 322
 QY 216 -----VOLVVEPEG-----GAVAP-----GCTVTLTCVTPAQSPQIHMMDGV 254
 Db 323 LDLDTMISLLSEPPQELLVNVSDVRVSPAAPERQEGSSLLTCEAESQJLLEFQWLKEET 382
 QY 255 PLPLPPSPVLLIPEIGPDQGYSCVAT-HSSHGPQESRAVYSIIEP 301
 Db 383 QGVLEKGPVLQHLHLKREAGGYKCVASVPSPGLNRTQLVNVALFPG 430

RESULT 5
 Q9BRD9

ID Q9BRD9 PRELIMINARY; PRT; 646 AA.
 AC Q9BRD9;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Similar to melanoma adhesion molecule.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 PP SEQUENCE FROM N.A.
 RA Straussberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC006329; AAH06329.1;
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 5.
 DR SMART: SM00408; IGc2; 3.
 DR SMART: SM00410; IG_Like; 1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 646 AA; 71608 MW; 1F2C696B64B16635 CRC64;

Query Match 12.48; Score 218.5; DB 4; Length 646;
 Best Local Similarity 23.38; Pred. No. 7.2e-09;
 Matches 95; Conservative 56; Mismatches 134; Indels 123; Gaps 19;

QY 4 ITARIGELPLVKYKGAPKKPP-QRLW--KLNTGRTEAWKVLSPG---GGPWDSVARVL 57
 Db 36 VEVEVGSTALLKCGLSQSGNLSHVDFSVHKKRKTLLIFRVGQGGQSEGEYEQRLSIQ 95
 QY 58 PNG-SLFLPAGVIGDEGIFRCQAMNRNKETKS-NYRV--RYVOIPGKPEI-----V 105
 Db 96 DRGATIALTVLPQDERIFLQ-----GKPPKQVEYRIQLPVYKAPPEPNQVNPGLGIPV 150
 QY 106 DSASELTAGVNVKVTGCVSGSYDAGTILSWHLDGKPLVPNEKGVSVKQTRRHPTGLFT 165
 Db 151 NSKE-----PEEVATCVGNGYPIQVIVYKNGRPLKEKNRVHI-QSSQTVSESSGLYT 203
 QY 166 LQSELMTVPARKSDPPPTFS--SFSP-ILPHFALP-----TAPI---QPPVW---EPVPL- 213
 Db 204 LQSLILKAQLVK-EDKDAQFYCELNYRLPSGNHMKESREVTVPFYPTKVMLEVEPVGML 262
 QY 214 -----EE----- 215
 Db 263 KEDRVEIKCLADGNPFPFHSISEUNPSTREAEETTNDUNGVLVLEPARKHESRYEQG 322
 QY 216 -----VOLVVEPEG-----GAVAP-----GCTVTLTCVTPAQSPQIHMMDGV 254
 Db 323 LDLDTMISLLSEPPQELLVNVSDVRVSPAAPERQEGSSLLTCEAESQJLLEFQWLKEET 382
 QY 255 PLPLPPSPVLLIPEIGPDQGYSCVAT-HSSHGPQESRAVYSIIEP 301
 Db 383 QGVLEKGPVLQHLHLKREAGGYKCVASVPSPGLNRTQLVNVALFPG 430

RESULT 6
 Q9Z214
 ID Q9Z214 PRELIMINARY; PRT; 1344 AA.
 AC Q9Z214;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Rlg-1 protein.
 GN RBIG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

DR	InterPro: IPR004006: Ig_MHC.
DR	Pfam: PF00041: Ig; 4.
DR	Pfam: PF00047: Ig; 5.
SC	SE-PIPE: 1534 AA, 15372 MW, 3013974.47 Da, 5664.
Query Match	
Best local similarity 12.18; Score 214; DB 4; Length 1934;	
Matches 77; Conservative 41; Mismatches 126; Gaps 12;	
QY	9 GEPVIGKCKGAPKKPPGRLKLNIGK...FAKKVLSQGGPPIESVAKVFNCS 61
DB	111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1

```

Brain Res. 640:296-307(1994)
CC 1- FUNCTION: CELL ADHESION MOLECULE THAT BINDS TO CD6. INVOLVED IN
CC NEURITE EXTENSION BY NEURONS VIA HETEROPHILIC AND HOMOPHILIC
CC INTERACTIONS. MAY PLAY A ROLE IN THE BINDING OF T AND B CELLS TO
CC ACTIVATED LEUKOCYTES, AS WELL AS IN INTERACTIONS BETWEEN CELLS OF
CC THE NERVOUS SYSTEM. MAY BE INVOLVED IN OSTEOGENESIS. MAY PLAY AN
CC IMPORTANT ROLE IN THE DEVELOPMENT OF A VARIETY OF MESENCHYMAL
CC TISSUES (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC 1- TISSUE SPECIFICITY: STRONGEST EXPRESSION IN THE LUNG, THEN BRAIN,
CC LIVER, AND KIDNEY. PRESENT IN THE SOMATOSENSORY SYSTEM, BASAL
CC GANGLIA, CORTEX, OLFACTORY SYSTEM, AND CIRCUMVENTRICULAR ORGANS.
CC 1- PTM: THE N-TERMINAL IS BLOCKED.
CC 1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 4
CC C2-LIKE AND 2 V-LIKE DOMAINS.
CC EXBL: AB004536; BAA23279.1;
CC EXBL: Y13241; CAA73693.1;
CC EXBL: Y13240; CAA73692.1;
CC HSP9: Q13740; IKJC.
CC InterPro: IPR003599; Ig_
CC InterPro: IPR003600; Ig_Like
CC InterPro: IPR003606; Ig_MHC.
CC Pfam: PF00047; Ig_5.
CC SMART: SM00409; Ig_3.
CC SMART: SM00410; Ig_Like_2.
CC Cell adhesion, immunoglobulin domain; Glycoprotein, Transmembrane;
KW Signal: Antigen.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 583 CD166 ANTIGEN.
FT DOMAIN 28 527 EXTRACELLULAR (POTENTIAL).
FT TRANSF 528 548 POTENTIAL.
FT DOMAIN 549 583 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 36 120 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 150 227 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 263 320 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 347 399 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 428 492 IG-LIKE C2-TYPE DOMAIN.
FT DISULF 43 113 POTENTIAL.
FT DISULF 157 220 POTENTIAL.
FT DISULF 270 313 POTENTIAL.
FT DISULF 354 392 POTENTIAL.
FT DISULF 435 485 POTENTIAL.
FT CONFLIC 339 339 S -> G (IN REF. 2).
FT CARBOHYD 95 95 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC...) (POTENTIAL).
CC SEQUENCE 583 AA; 65021 MW; D36B73854F5DF61E CRC64.

Query Match 12.1%; Score 213; DB 11; Length 583;
Best Local Similarity 25.3%; Pred. No. 1.7e-08;
Matches 82; Conservative 50; Mismatches 136; Indels 56; Gaps 13;

QY 4 ITARIEPIVLKCKGAPKPPQPLEWKLNTGRTFAKVLSPQGGWDSVAVPLPNCISLFLPAV- 67
DB 31 VNSAYGDTIWMPPR---IIVPQNIIMF-----GK---WKYFKPKTSWFIAPFSTKKSVDY 80
QY 51 DVSAR-----VLPNGSFLIPAVGICQDGIIFRCQAMNRCNCKTKSNY-----RVRYQIP 99
DB 81 DQVPEYKDRLSLSENLTSLNNAKISDEKPEFVCMV-----TEDVVEAPTIVKVKQPF 134
QY 100 GKPEIVDASLILAGVPNKVGICVSGYPAGTISWHLDDKPLIVPNEKSVSKQETRRHP 159
DB 135 SKPEIVNRAAPFTDQIKLGLTGLTSMISVPHQNIWYFNGKVLQVVDGEVSLFKKEILIP 194
QY 160 ETSGLFTQLSMLMTPARSGDPPPTFSQS---FSPGLPRLRALPTAPIQPPVWE-PVPLEE 215
DB 195 GTQLYNTSSLSLVKTKK-SLIQMLPFTCSVIYVYGPS-----GUKTIYSPQALDIYVYPRQ 248

```

```

QY 216 VOLVVEPCGAVAVACCIPTLTCEVPAQVLSK QIHWKOCVPLPLPSVLLIFIGHQDQ 274
DB 249 VTIQVLPKNAIKESDNITLQCLNGNPPPEFMYLPQAGIRSSNTYTLTDVRRNT 308
QY 275 GTYSVATHSSHGQSPAVSISI 298
DB 309 GDYKC-----SLIDORRNAASTTI 327

RESULT 9
QY0242
ID QY0242 PRELIMINARY; PRT; 330 AA.
AC QY0242;
BT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
ST 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Roundabout1 protein (Fragment).
GN ROBO1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
RX MEG...NI-21356016; PubMed 11472852;
RA Vargesson N, Iuria V, Messina I, Erskine L, Laufer E.;
RT "Expression patterns of Slit and Robo family members during vertebrate
RT limb development."
PL Mech. Dev. 106:175-190(2001).
DR EMBL; AF364047; AAK94293.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00406; Ig_v; 1.
FT NON_TER 1
FT NON_TER 330
FT NON_TER 330
CC SEQUENCE 330 AA; 36745 MW; 0613488F78CEBE61 CRC64;

Query Match 11.9%; Score 210.5; DB 13; Length 330;
Best Local Similarity 24.7%; Pred. No. 1.3e-08;
Matches 76; Conservative 45; Mismatches 136; Indels 51; Gaps 11;

QY 9 GPEVLKCKGAPKPPQPLEWKLNTGRTFAKVLSPQGGWDSVAVPLPNCISLFLPAV- 67
DB 37 GEPATLCK-AEGRFT-TIENYKGERVETIKQ-DPK-----SH-MLLPSSSLEFLRIV 88
QY 68 -----GIQDGIIFRCQAMNRCNCKTKSNYRVRYQIPGKPEIVDASLILAGVFN-KVGTG 122
DB 89 HGKRSRDEGVVVCVARNYLGDAVSHNASLEVAII--RDDFRQNPDSVMVAVGPPAVMEC 146
QY 123 VSENSYPATISWHITGTFPLVPNEKSVSKVETKPHPETGLTLQSEIMVTPARGSDPPP 182
DB 147 QPPRGHPPTISWKKDSTPLDQKDEIFIR-----GQKLMIITRKND--- 189
QY 183 FSCSEFSLPLPHHAI-TAPIQPR---VWEKVPLEEVULVVEPCGAVACCIPTLTCEV 239
DB 190 -----AKYVVCVGTNNVGERSEVAELITVLEPPSFVKRPSN-LAVTVDSDREFKCEA 240
QY 240 PAQSPSQIHWKOCVPLPLPSVLLIFIGHQDQ-----LPSVWLIIPETICQDQGTYSV-ATHSSHGQSPAV 294
DB 241 RQDPVPTVWKKKQDGEIDPKARYEIRHDIHTLTKIKVMAGIMSGSYTCVAHMMVGKAKASATL 300
QY 295 SISIIEP 301
DB 301 IVQVVSEP 308

RESULT 10
QY0255
ID QY0255 PRELIMINARY; PRT; 149 AA.
AC QY0255;
DT 01-FEB-1997 (TREMBLrel. 02, Created)

```

DT 01 FEB 1997 (Tremblrel. 02, last sequence update)
 DE 01 JUN 2002 (Tremblrel. 21, last annotation update)
 DE MYELOBLAST KIAA0230 (Fragment)
 GN KIAA0240.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cladiala; Vertebrata; Euteleostomi.
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A.
 RC TISSUE= BONE MARROW;
 RA MEDLINE=97191544; PubMed 9099022;
 RA Napsse T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
 RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RI "Prediction of the coding sequences of unidentified human genes. VI.
 RI The coding sequences of an new genes (KIAA0201-KIAA0280) deduced by
 RI analysis of cDNA clones from cell line KG-1 and brain.";
 RL DNA Res. 3: 321-329(1996).
 RN 121
 RP SEQUENCE FROM N.A.
 RC TISSUE= MELANOMA;
 RA MEDLINE=95048383; PubMed 7959781;
 RA Weiler S.R., Taylor S.M., Evans R.J., Kan Mitchell J., Mitchell M.S.,
 RA Trent J.M.;
 RI "Assignment of a human melanoma associated gene M650 (M2S4448) to
 RI chromosome 2p25.3 by fluorescence in situ hybridization.";
 RI Genomics 22: 243-244(1994).
 RN 131
 RP SEQUENCE FROM N.A.
 RC TISSUE= MELANOMA;
 RA Mitchell M.S., Kan Mitchell J., Minev R., Edman G., Evans R.J.;
 RI "Identification of a novel melanoma gene (M650) - likely the gene for
 RI 111 receptor antagonist, which encodes epitopes recognized by human
 RI cytolytic T lymphocytes.";
 RI Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: D86984; BAA13219.1;
 DR EMBL: AF200448; AAT06454.1;
 DR HSSP: P05164; 1CXP.
 DR InterPro: IPR002007; Anm.peroxidase.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003591; LRR_Typ.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00047; Iq_4.
 DR Pfam: PF00560; LRR_5.
 DR Pfam: PF01464; LRRCT_1.
 DR Pfam: PF00094; VWC_1.
 DR PRINTS: PR00457; ANTEROXIDASE.
 DR SMART: SM00408; IGG2_4.
 DR SMART: SM00082; IPRCT_1.
 DR SMART: SM00013; IPRNT_1.
 DR SMART: SM00469; LRR_Typ_4.
 DR SMART: SM00214; VWC_1.
 DR PROSITE: PS01208; VWF; UNKNOWN 1.
 KW Immunoglobulin domain.
 FT NON TER
 SQ SEQUENCE 1496 AA; 167209 MW; E909A7069B1ABEF CRC64;
 Query Match 11.8%; Score 209; DB 11; Length 1612;
 Best Local Similarity 25.4%; Pred. No. 146 07;
 Matches 78; Conservative 44; Mismatches 146; Indels 52; Gaps 11;
 QY 2 QNITAPGFEIVIKKSAFKKFKLQFEWELNIGELAEKVEVLSQKQKWSVAPVFNCS 61
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 466 ENTEVLGVSVELR SAUGHGPRISW-TRGRIHL----- PVPVNVNIPSSG 413
 QY 62 LPTPAVCIQDQCIKPRGAMNNKSKPKSNYPVYVCIIPKPEIVASVETATVWNVN 121
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 414 LYTQNVQGSGEYA/SATNINDSVIATAF-----ITVALPQVTPQPRV-- 460

QY 122 CYSESSYPATILSWLLDGR PLVNEKGVSKQIERRRSTGLHLSHLMVTPAR 177
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 461 -VIEGG TVDFQCEAKGNPPVIAWTKGSGQSVDPRIHVLSSSTHRSVVAHQQG 515
 QY 178 GSKPKTESSESPGPRHAKLKAIPQPRVMEVPIEEVQAVVEFGAVARGLVITTC 247
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 516 ---QYECQAVNIIQSGQKVVAILTVQPRV-TPV FASIDSTTVGVCANVLIC 564
 QY 238 EYPAQISPTIHWKIKCVPLP LIPSPWLLIPETIGPQIGYSRVAHSHSHDIPS 291
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 564 SSQREFFPATWTKKIKVAVTESKRIKSPESLITINWIPALAVPEVAINTIGASVS 624
 QY 292 RAVSISLIEEGEG 405
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 624 MVLISVNVVIVSRNG 637
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 11
 ID 089026 PRELIMINARY; PRI: 1612 AA.
 AC 089026;
 DT 01 NOV-1998 (Tremblrel. 08, Created)
 DT 01 NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01 DEC-2001 (Tremblrel. 19, Last annotation update)
 DE DUT11 protein.
 GN RHOBL OR DUT11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 111
 RP SEQUENCE FROM N.A.
 RC TISSUE= BRAIN;
 RA Wu M.C., Lowe N., Fordham R., Rabbitts P.;
 RI "The mouse homologue of human DUT11/RHOBL gene: protein sequence and
 RI chromosomal location.";
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: V17793; CAA76850.1;
 DR HSSP: P56276; IRLK.
 DR MCD: MCL1274781; RhoBL.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003600; Iq_Like.
 DR InterPro: IPR003006; Iq_Like.
 DR Pfam: PF00041; fn3_4.
 DR Pfam: PF00047; Iq_5.
 DR SMART: SM00060; FN3_4.
 DR SMART: SM00406; Iq_1.
 DR SMART: SM00410; Iq_IIIb; 4.
 SQ SEQUENCE 1612 AA; 176406 MW; 5f2988c54479648b CRC64;
 Query Match 11.8%; Score 209; DB 11; Length 1612;
 Best Local Similarity 25.4%; Pred. No. 146 07;
 Matches 78; Conservative 44; Mismatches 146; Indels 52; Gaps 12;
 QY 9 GELVILKGAIKKIKQPIELKWKUNIGRIEAKKVISSQGGDWSVAVRIPNKSIFAV 67
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 43 GEPATLNC AKGRPIPIEHWKGRVLEKLD DGR SHRMILPSGSELEIRIV 94
 QY 68 ---GIQPPGPIPGVAMNPKSKPKSNYPVYVCIIPKPEIVASVETATVWNVN 122
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 95 HGKSKSGQGVTCVARNYLDAVSHNASTVALIL KQDFKQINIVVMVAVGDAVMB 152
 QY 123 VSEGSYPATILSWLLDGR PLVNEKGVSKQIERRRSTGLHLSHLMVTPAR 182
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 103 QFERCHPEPTISWKKKSGSELGKQKERTIR GSKMLITYIKSD 195
 QY 183 IFCSFSFGLRHKALRTAPQPR---VWEIVPIEEVQAVVEFGAVARGLVITTC 249
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 196 -----ACKYVCGVNMWTERSHVAELIVIRTSFVKRGLAVTVGSAIFKKA 246
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 240 PAQSPQIHWKIKCVPLP-----LPPSPVILPEIGPQDQCIYSKVAHSHSHDIPS 294

Query Match
Best Local Similarity 11.8%; Score 308; DB 5; Length 1033;
Matches 88; Conservative 58; Mismatches 139; Indels 108; Gaps 18;
QY 12 LVLKCG--KGAD-KKPPQRLKWLKNTGRTEAKVLSPOGGPWSVARVLPNGSLFLPAVG 68
DB 133 LLLKCHVEGASGDLPLEIEWRNSEKLSWKVQ-----LDGHLITIPQ 179
QY 69 IODEGIFPCOAMNRNGK-ETKSNYRVRYQ-----IPCKP----- 102
DB 180 SEDDGLYKCTASNAAGRVMSKQGY---VQSSVKCLPRIPPRKQKMMESWIKQTFICRG 236
QY 103 -----EIVDSASELTAGVPNKVG-TCVSEGSYPAGT-----LSWILD 138
DB 237 KRGGCAAGLEALPAAPEDRLIVQGGPIGQSIIKEGHTALTCLYELPFLKNCPIQLRWKK 296
QY 139 CKPLVPNKGVN-----VKEQTR--RHPETGIFTIQSLMTVPARGGDP 181
DB 297 GKLLPQVLSGASPPGHSFGSKDALLPEIAPLVHLKQNG--TLSPASIIASDAG---- 350
QY 182 PTFCSSFGPLPKRHRLTAPLOPWEPVLEFVQVVEPEGGAAPGGTIVT-ITCEVP 240
DB 351 -OYQCQLQ--LEAHAPINSSPGILEV-----IEQLKFPQPTSKNLELDVAVKVCRAQ 402
QY 241 AUFSEFQIHWKCKVPLPLP-----PSPVLILPELSPQGGTYSVATHSSHSQESRAV 294
DB 403 GTPTFQVCVVRDSENTLLPDHVEVDANGTLIFRNVNSBRGNYTCLATNSQGINATVAI 462
QY 295 SISIIIEPCEEGPTAGSVGGSGLTALALGILG 327
DB 463 NV-VVTPKFSVPVPGFIETSEQGVVVMHCQAIG 494

Q24327 PRELIMINARY: PPT; 1033 AA.
AC Q24327;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE GP160-DTRK precursor
GN CTK OP DTRK OR CGR067.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
PP SEQUENCE FROM N.A.
RC STRAIN=CANTON S;
RX MEDLINE=92164624; PubMed=1371458;
RA Pulido D., Campuzano S., Koda T., Modolell J., Barbacid M.;
KA "Dlrc, a Drosophila gene related to the trk family of neurotrophin
RT receptors, encodes a novel class of neural cell adhesion molecule.";
PL EMBO J. 11:391-404 (1992)
NP EMBL; X63453; CAA45053.1;
DR HSSP; P12931; IPWK.
DR Flybase; FBgn0004839; otk.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Iq_C2.
DR InterPro; IPR003600; Iq_Like.
DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Iq_5.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00408; Iqez; 3.
DR SMART; SM00410; Iq_Like; 2.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00117; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00130; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Cell adhesion; Immunoglobulin domain; Signal;
KW Transferase; Tyrosine-protein kinase.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 1033 Gp160-DTRK.
SQ SEQUENCE 1033 AA; 114367 MW; D282EFCB28ACA8D0 CRC64;

Query Match
Best Local Similarity 25.4%; Score 289; DB 4; Length 1451;
Matches 78; Conservative 44; Mismatches 133; Indels 52; Gaps 12;
QY 9 GEPILVLKCKCAPKKPQRLKWLKNTGRTEAKVLSPOGGPWSVARVLPNGSLFLPAV- 67
DB 82 GEPATLNCK-AGGRPTPIIEWYKGGKVEIDKD-DPR-----SIRMLLPSSGSLFLRIV 133
QY 68 -----GIDEGIFPCOAMNRNGK-ETKSNYRVRYQIPCKPEIVDSASELTAGVFN-KVGTC 122
DB 134 HGKKSRLPDEGVYCVVAKNVLGAVSHNASLEVAAIL--RDDFRQNP-SDVMVAVGEFAVMBC 191
QY 123 VSEGSYPAGTFLSWILDQKFLVNEKGVSVKELTKRHPETHLFTLASELMTVPARGGSDPPP 182
DB 192 QPPKCHPEPIISWKKKCSPLLUKDFEITIP-----GCKIM:TVTRKSD--- 234
QY 183 TPSCSFGPLPKRHRLTAPLOPWEPVLEFVQVVEPEGGAAPGGTIVTITCEV 239
DB 235 -----ASKYVWV:INNVGSEPESEVAELTVLEPPESEVKRNSNLAVTVDDSAEPKCEA 285
QY 240 PAQSPQIHWKCKVPLP-----LPIVSPVLILPEIGPQGGTYSVATHSSHSQESRAV 294
DB 286 RGDVPTVTRKIDVGLPKSPYRIRDDHTLIKIRKVTAGDMGVSVCVA-ENMVGKADASA- 343
QY 295 SISIIIEP 301
DB 344 TLTVORP 350

RESULT 13
Q24327

Fri May 30 17:16:40 2003

Search completed: May 30, 2003, 15:57:30
Job time : 79.018 secs

us-09-872-185b-2.rspt

Page 9

